

Phonological and Physiological Constraints on Assimilatory Pharyngealization in Arabic: Ultrasound Study

Sawsan Alwabari

A thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
Doctorate in Philosophy degree in Linguistics

Department of Linguistics
Faculty of Arts
University of Ottawa



Dedication

For my parents, who taught me the first sounds of language, and for my lovely husband, who patiently supported me throughout my doctoral journey.

Acknowledgments

I am immensely grateful to my supervisors, Prof. Marc Brunelle and Prof. Jeff Mielke, for their commitment to my academic success. I am thankful to them for dedicating their valuable time to my research and for patiently guiding me through this learning process. I always believe that I am fortunate for having Marc and Jeff as my supervisors, and the compliments I received at various conferences when talking about my supervisors and my research made me feel that it is an honor to work with them. I would never have dreamed of a better, more informative supervision experience than my experience with Marc and Jeff. ‘Thank you’ is never enough.

I extend my thanks to Prof. Tania Zamuner for supervising my previous research in psycholinguistics and for giving me the opportunity to learn how to collect and analyze eyetracking data. Thank you for her support and for this astonishing experience.

I am also grateful to the members of my thesis committee, Khalil Iskarous, Christopher Fennel, Kevin McMullin and Tania Zamuner. I acknowledge their constructive feedback and insightful questions that have contributed to enriching this dissertation.

I extend my profound gratitude to Prof. Alexei Kochetov for providing me the opportunity to collect a major portion of my ultrasound data at his Phonetics Lab, University of Toronto. I cannot but appreciate the generosity and welcoming attitude of Alexei and his lab members who allocated a great deal of the busy lab schedule for my experiment. Their cooperation allowed me to collect the data in a remarkably short amount of time. Thank you, Alexei.

I have been lucky to be surrounded by great friends in the department of linguistics who have been always eager to lend a helping hand. Their unwavering support has always helped me in relieving the stress associated with completing a doctorate degree. Thanks to Lyra Magloughlin for always giving me a sense of support while performing this research. Thanks to Jonathan Lim

for giving me valuable advice on my first day in the PhD program and for always being willing to help. Thanks to Jumanah Abusulaiman, Vesela Simeonova, Kiranpreet Nara, Reza Falahati and Tan Ta for making my journey full of energy and support.

This dissertation would never have existed without the invaluable contribution of my research participants and recruitment coordinators in Ottawa and Toronto. I thank them all for their time and generosity. I also thank Maurice Bélanger, the department's genius IT who was always present to assist me in running my experiments at the Sound Pattern Laboratory.

I cannot express my sincere gratitude to my family. I am greatly indebted to my parents, siblings, husband and mother-in-law – words can never do them justice. Their compassion and love have always been a source of relief for me. My parents had to endure my absence for ten years and, I know, this is an enormous sacrifice that they gave so that I can achieve my goals. I am profoundly grateful to them for their unconditional love and continuous support. My lovely husband deserves more than 'thank you' for his patience and encouragement over the past years.

I love you!

Table of Contents

List of Figures	xii
List of Tables	xix
Abstract	xx
CHAPTER 1 Physiology of Speech and Pharyngealization	1
1.1 Introduction.....	1
1.2 Anatomy of the Speech Apparatus and Speech Physiology	5
1.2.1 The tongue	5
1.2.1.1 Extrinsic muscles.....	6
1.2.1.2 Intrinsic muscles.....	9
1.2.2 The pharynx.....	11
1.3 Emphasis.....	13
1.3.1 Articulation of emphasis.....	13
1.3.1.1 Articulation of uvulars and pharyngeals	13
1.3.1.2 Articulation of emphatics	17
1.3.1.2.1 Emphasis as uvularization.....	17
1.3.1.2.2 Emphasis as pharyngealization	19
1.3.1.2.3 Variation in emphasis articulation	26
1.3.2 Acoustics of emphasis	29
1.3.2.1 Acoustic evidence for pharyngealized or uvularized constrictions.....	30
1.3.2.2 Other factors affecting the properties of coarticulated vowels.....	31
1.4 Dialect Typology	33
1.5 Structure of the Dissertation	35

CHAPTER 2 Coarticulation: Theories and Predictions.....	38
2.1 Introduction.....	38
2.2 Assimilation as Feature Spreading and Coarticulation as Coproduction	39
2.3 Factors Affecting Coarticulation	42
2.3.1 Phonological contrast	42
2.3.1.1 Manuel's hypothesis about phonological contrast	42
2.3.1.2 Limitations of the contrast hypothesis.....	44
2.3.1.3 Applying the contrast hypothesis to other coarticulation phenomena	46
2.3.2 Physiological constraints	48
2.3.2.1 Defining gestural conflict.....	50
2.3.2.2 Coarticulation and strategies for solving gestural conflict.....	51
2.3.2.2.1 Multiplicity of reconciliation strategies	52
2.3.2.2.2 Resolving inter-articulator conflict	54
2.4 Coarticulation Models	56
2.4.1 Feature-based models	57
2.4.1.1 Early feature-based models	57
2.4.1.2 Recent feature-based models.....	58
2.4.2 Gesture-based models.....	63
2.4.2.1 Early models of coarticulation resistance.....	63
2.4.2.2 The degree of articulatory constraints (DAC) model.....	65
2.4.2.3 Comparison of coarticulation resistance models and preference for DAC model	69
2.5 Predictions of Emphasis Spread	71
2.5.1 Predicted effects of phonological constraints and the feature-based model.....	72
2.5.1.1 Predictions based on phonological contrast	74

2.5.1.2 Predictions of the feature-based model	74
2.5.2 Predicted effects of physiological constraints and gesture-based model	76
2.5.2.1 Predictions based on physiological constraints	77
2.5.2.2 Predictions of a gesture-based model	79
CHAPTER 3 Is Emphasis Spread Categorical?	82
3.1 Introduction	82
3.2 Research Questions and Hypotheses	83
3.2.1 Uvularization vs. pharyngealization	83
3.2.2 Categorical vs. gradient emphasis	84
3.3 Methodology	85
3.3.1 Participants	86
3.3.2 Stimuli	86
3.3.2.1 Stimuli conditions	86
3.3.2.2 Stimuli selection criteria: Minimizing noise	89
3.3.2.2.1 Minimizing phonetic and phonological variation	89
3.3.2.2.2 Minimizing prosodic variation	90
3.3.2.2.3 Minimizing lexical variation	91
3.3.3 Data collection	91
3.3.4 Data processing	93
3.3.5 Data analysis	95
3.3.5.1 Using polar versus Cartesian SS-ANOVA	97
3.3.5.2 Limitations of SS-ANOVA	98
3.3.5.3 Additional analyses	100
3.3.5.3.1 Scatterplots for quantifying differences of tongue fits	100

3.3.5.3.2 Retraction ratios for the size of coarticulation relative to plain and emphatic references.....	103
3.3.5.3.3 ANOVA for categorizing coarticulation and assimilation.....	104
3.4 Results.....	108
3.4.1 Uvularization vs. pharyngealization (Question 1).....	109
3.4.2 Categorical vs. gradient assimilation (Question 2).....	114
3.4.2.1 Assimilation in /s/.....	115
3.4.2.2 Assimilation in /t/	119
3.4.2.3 Tongue root and body differences in /s t/ assimilation by speaker	123
3.5 Discussion.....	125
3.5.1 Lingual configuration in phonemic pharyngealization.....	125
3.5.2 Categoricalness of assimilatory pharyngealization	126
CHAPTER 4 Phonological Factors and Assimilation	128
4.1 Introduction.....	128
4.2 Research Question and Hypothesis	129
4.3 Methodology	130
4.4 Results.....	132
4.4.1 Assimilatory pharyngealization in individual consonants.....	133
4.4.1.1 Contrastive coronals /t s/	133
4.4.1.2 Non-contrastive coronals /n l/	134
4.4.1.2.1 coarticulation in /n/	134
4.4.1.2.2 Assimilation in /l/	137
4.4.1.2.3 Tongue root and body difference in /n l/ assimilation by speaker ..	142
4.4.2 Comparison of assimilation between contrastive and non-contrastive coronals. 144	
4.4.2.1 Coronals in plain context.....	145

4.4.2.2 Coronals in pharyngealized context	147
4.4.2.3 Magnitude of assimilation in all coronals	150
4.5 Discussion.....	153
CHAPTER 5 Physiological Constraints on Coarticulation	158
5.1 Introduction.....	158
5.2 Research Question and Hypotheses.....	159
5.2.1 Compromise hypothesis	159
5.2.2 Resistance hypothesis	160
5.3 Methodology.....	161
5.4 Results.....	164
5.4.1 Labials	167
5.4.2 Palatals.....	174
5.4.3 Velars.....	180
5.4.3.1 Resistance in [g]	182
5.4.3.2 Resistance and compromise in [χ].....	185
5.4.4 Comparison of all target consonants	189
5.5 Discussion.....	192
5.5.1 Gestural antagonism and phonemic pharyngealization	192
5.5.2 Coarticulatory tongue root retraction	193
CHAPTER 6 General Discussion	195
6.1 Introduction.....	195
6.2 Lingual Configuration in Producing Emphatics	196
6.3 Constraints on Assimilatory Pharyngealization.....	198
6.3.1 Phonological constraints on assimilatory pharyngealization	198
6.3.2 Physiological constraints on assimilatory pharyngealization.....	200

6.4 Phonetics and Phonology of Assimilatory Pharyngealization	204
6.5 Pharyngealization and Models of Coarticulation.....	208
6.5.1 Pharyngealization and the target-interpolation model.....	208
6.5.1.1 Results raising concerns about the model	209
6.5.1.2 Evidence supporting the model	210
6.5.2 Pharyngealization results and the DAC model.....	213
6.6 Future Directions	215
6.7 Conclusion	217
References.....	219
Appendices.....	231

List of Figures

<i>Figure 1.1. The speech apparatus showing the cervical vertebrae and extrinsic lingual muscles: GG= genioglossus, HG= Hyoglossus, PG= palatoglossus and SG= styloglossus. GH= geniohyoid (a non-lingual muscle). Reproduced from Crumbie, Salvador, and Rad (2019).</i>	<i>7</i>
<i>Figure 1.2. A coronal view of the tongue illustrating the intrinsic and some extrinsic lingual muscles. Reproduced from Crumbie et al. (2019).....</i>	<i>10</i>
<i>Figure 1.3. Sagittal (left) and posterior (right) view of the speech apparatus showing the superior, middle and inferior pharyngeal constrictors.</i>	<i>12</i>
<i>Figure 1.4. Vocal tract configuration in producing pharyngeals /ħ ʕ/ showing pharyngeal compression. Reproduced from Lawson et al. (2018).</i>	<i>15</i>
<i>Figure 1.5. Vocal tract configuration in producing uvular /ʁ/ (left) and /q/ (right) showing uvular aperture and no epiglottal retraction. Reproduced from Lawson et al. (2018).</i>	<i>17</i>
<i>Figure 1.6. Vocal tract configuration in producing emphatics as uvularized showing the secondary articulation at the uvula. Reproduced from Bin-Muqbil (2006).....</i>	<i>19</i>
<i>Figure 3.1. Alignment of audio recordings by locating the stop release on the external audio (top) and on Ultraspeech audio (bottom). An interval of length 56.842982 (top) minus 0.268651 (bottom) was removed from the start of the external audio.</i>	<i>94</i>
<i>Figure 3.2. Ultrasound image illustrating the measurement of the occlusal plane angle per speaker as indicated by the green line.....</i>	<i>96</i>
<i>Figure 3.3. Original tongue traces from speaker S06 showing (A) original un-rotated splines and (B) splines rotated to 14.25°.</i>	<i>97</i>
<i>Figure 3.4. SS-ANOVA using Cartesian (left) and polar coordinates (right), reproduced from Mielke (2015).....</i>	<i>98</i>

Figure 3.5. SS-ANOVA fit (tongue tip to the right) and 95% CI of three example conditions.	101
Figure 3.6. The difference between SS-ANOVA fits of the three conditions shown in the sample SS-ANOVA for all speakers. A negative root difference in the left means Coarticulated is less retracted than Emphatic and a positive root difference in the right means Coarticulated is more retracted than Plain.	103
Figure 3.7. Example showing a significant, yet inconspicuous, difference between coarticulated and distinctive /s/'s.	106
Figure 3.8. Ultrasound images of speaker S17 showing the tongue shape at the release of /t ^ɕ / in /t ^ɕ a.fi:f/ 'shallow' (left) and /t/ in /tatli:f/ 'ruining' (right).	109
Figure 3.9. SS-ANOVA fit at mid-phone/closure (tongue tip to the right) and 95% CI showing C ^ɕ in plain and in emphatic context and plain C.	110
Figure 3.10. SS-ANOVA fit and 95% CI showing distinctive C ^ɕ at all selected time points.	111
Figure 3.11. The difference between SS-ANOVA fits at mid-phone/closure for all speakers: distinctive C ^ɕ in plain context relative to the same consonant in pharyngealized context (left) and to plain C (right) where the negative root difference means more retraction in distinctive C ^ɕ ...	112
Figure 3.12. SS-ANOVA fit at 50% of the phone/closure and 95% CI showing /t s t ^ɕ s ^ɕ / in plain context.	113
Figure 3.13. The difference between SS-ANOVA fits of /t ^ɕ / minus /s ^ɕ / in plain context for all speakers. Positive root difference and negative body difference mean that the root in /s ^ɕ / is more retracted and the body is higher than in /t ^ɕ /.	114
Figure 3.14. SS-ANOVA fit at mid-phone and 95% CI showing categorical assimilation in /s/.	116

Figure 3.15. ANOVA and root retraction ratio of mean root radius showing categorical assimilation in /s/.	117
Figure 3.16. SS-ANOVA fit showing assimilated /s/ at three selected time points.	118
Figure 3.17. SS-ANOVA fit at mid-closure and 95% CI showing resistance in /t/.	119
Figure 3.18. ANOVA and root retraction ratio of mean root radius showing resistance in /t/... ..	120
Figure 3.19. SS-ANOVA fit at mid-closure and 95% CI showing assimilation in /t/.	121
Figure 3.20. ANOVA and root retraction ratio of mean root radius showing categorical assimilation in /t/.	121
Figure 3.21. SS-ANOVA fit showing assimilated /t/ at two selected time points.	122
Figure 3.22. The difference between SS-ANOVA fits comparing assimilated /t s/ to distinctive /t ^s s ^s / (left) and to plain /t s/ (right) for all speakers. Negative root difference in the left means the assimilated C is less retracted than distinctive C ^s , and positive difference in the right means assimilated C is more retracted than plain C.	124
Figure 4.1. The difference between SS-ANOVA fits comparing assimilated /t s/ to distinctive /t ^s s ^s / and to plain to for all speakers.	134
Figure 4.2. SS-ANOVA fit at mid-phone and 95% CI showing resistance in /n/.	135
Figure 4.3. ANOVA and root retraction ratio of mean root radius showing resistance in /n/. ...	136
Figure 4.4. SS-ANOVA fit of /n/ in pharyngealized context at three selected time points.	137
Figure 4.5. SS-ANOVA fit at mid-phone and 95% CI showing categorical assimilation in /l/. ..	138
Figure 4.6. ANOVA and root retraction ratio of mean root radius showing categorical assimilation in /l/.	139
Figure 4.7. SS-ANOVA fit at mid-phone and 95% CI showing gradient coarticulation in /l/. ..	140

Figure 4.8. ANOVA and root retraction ratio of mean root radius showing gradient coarticulation in /l/.	141
Figure 4.9. SS-ANOVA fit of /l/ in pharyngealized context at three selected time points.....	142
Figure 4.10. The difference between SS-ANOVA fits comparing assimilated /n l/ to distinctive C ^ɕ (left) and to plain /n l/ (right) for all speakers. Negative root difference in the left means the assimilated C is less retracted than distinctive C ^ɕ , and positive difference in the right means assimilated C is more retracted than plain C.	143
Figure 4.11. SS-ANOVA fit at mid-phone/closure and 95% CI showing (non)-contrastive coronals in plain context.	146
Figure 4.12. The difference between SS-ANOVA fits comparing each two coronals in plain context for every speaker. Negative root difference in the left means /s/ was less retracted than /l/ and in the right means /t/ was less retracted than /n/.	147
Figure 4.13. SS-ANOVA fit at mid-phone/closure and 95% CI showing (non)-contrastive coronals in pharyngealized context.....	148
Figure 4.14. The difference between SS-ANOVA fits comparing each two coronals in pharyngealized context for every speaker. Negative root difference in the left means /s/ was less retracted than /l/ and in the right means /t/ was less retracted than /n/.	149
Figure 4.15. The difference between SS-ANOVA fits of all coronals comparing plain vs. assimilated C for every speaker. Positive root difference means that assimilated C was more retracted than plain C	151
Figure 4.16. Root retraction ratio for the degree of assimilatory root retraction in /s l n t/, showing more retraction in /s/ than /l/ and in /t/ than /n/.	152

Figure 4.17. Root retraction ratio for the degree of assimilatory root retraction in /s l n t/, showing less retraction in /s/ than /l/ (although both were categorically assimilated).	152
Figure 5.1. SS-ANOVA fit at mid-phone/closure and 95% CI showing plain and pharyngealized C2.....	165
Figure 5.2. SS-ANOVA fit and 95% CI showing C ^s at all selected time points.	166
Figure 5.3. SS-ANOVA fit at mid-phone/closure and 95% CI showing all C2s /t t ^s s s ^s /.	166
Figure 5.4. SS-ANOVA fit at mid-phone/closure and 95% CI showing C ^s preceded by labials, palatals and velars.	167
Figure 5.5. Ultrasound images from speaker S17 at mid-closure of labial [b] in plain context (left) and pharyngealized context (right) from the same pair /basamha: - bas ^s amha:/.	168
Figure 5.6. SS-ANOVA fit at mid-phone and 95% CI showing coarticulation in labials.	169
Figure 5.7. SS-ANOVA fit at mid-phone and 95% CI showing assimilation in labials.....	169
Figure 5.8. The difference between SS-ANOVA fits of coarticulated labials vs plain ones (left) and vs distinctive C ^s (right) for all speakers. Positive root difference in the left means that the coarticulated labials are more retracted than plain ones.	171
Figure 5.9. ANOVA and ratio of mean root radius showing coarticulation in labials.	172
Figure 5.10. SS-ANOVA fit at mid-phone and 95% CI showing coarticulation in [b] vs [f]. ...	173
Figure 5.11. SS-ANOVA fit of labials in pharyngealized context at all selected time points....	174
Figure 5.12. SS-ANOVA fit at mid-phone and 95% CI showing palatals [ʃ j] in both contexts.	175
Figure 5.13. SS-ANOVA fit at mid-phone and 95% CI showing palatal [ʃ] vs. [j] in both contexts.	176
Figure 5.14. The difference between SS-ANOVA fits for individual palatals [j ʃ] in pharyngealized vs plain context (left) and vs source C ^s (right) for all speakers. Negative root and	

body difference in the right means that the palatals had less retracted root and higher body than source C ^ʕ	177
Figure 5.15. ANOVA and ratio of mean root radius of palatal [j] and [ʃ] in both contexts.	178
Figure 5.16. SS-ANOVA fit showing [j] in pharyngealized context at all selected time points.	179
Figure 5.17. SS-ANOVA fit showing [ʃ] in pharyngealized context at all selected time points.	179
Figure 5.18. SS-ANOVA fit at mid-phone and 95% CI of [χ] vs. [g] in plain context, with [χ] retraction.	181
Figure 5.19. SS-ANOVA fit at mid-phone and 95% CI of [χ] vs. [g] in plain context with no [χ] retraction.	181
Figure 5.20. SS-ANOVA fit at mid-closure and 95% CI showing resistance velar [g].	182
Figure 5.21. The difference between SS-ANOVA fits for individual velars [χ g] in pharyngealized vs plain context (left) and vs source C ^ʕ (right) for all speakers. Negative root and body difference in the right means that the velar had less retracted root and higher body than source C ^ʕ	183
Figure 5.22. ANOVA and ratio of mean root radius showing velar [g] in both contexts.....	184
Figure 5.23. SS-ANOVA fit of [g] in pharyngealized context at two selected time points.	185
Figure 5.24. SS-ANOVA fit at mid-closure and 95% CI showing velar [χ] slight dorsal lowering in pharyngealized context and distant root position from C ^ʕ	186
Figure 5.25. SS-ANOVA fit at mid-closure and 95% CI showing resistance of velar [χ] and its proximity to the root position of C ^ʕ	186
Figure 5.26. ANOVA and ratio of mean root radius showing uvular [χ] in both contexts showing higher retraction ratio before C ^ʕ	188

Figure 5.27. ANOVA and ratio of mean root radius showing uvular [χ] in both contexts showing similar retraction ratio before C ^ɕ and C.	189
Figure 5.28. SS-ANOVA fit at mid-closure and 95% CI showing labials, palatals and velar/uvular in plain context.	191
Figure 5.29. SS-ANOVA fit at mid-closure and 95% CI showing labials, palatals and velar/uvular in pharyngealized context.	192

List of Tables

Table 1.1 Comparison of the articulation of uvulars, pharyngeals and emphatics	23
Table 1.2 Survey of studies supporting uvularization and pharyngealization	27
Table 1.3 Chart of the Arabic consonant inventory, adapted from Thelwall and Sa'adeddin (1999).....	34
Table 2.1 Summary of predicted effects of pharyngealization on all consonants according to the type of constraints on the consonant	71
Table 3.1 The complete wordlist used in Experiment 1 showing the four test conditions (stress in boldface)	87
Table 4.1 Complete list of words used to address the effect of phoneme contrast	131
Table 4.2 Assimilatory effects (by consonant) in pharyngealized context showing the number of speakers	145
Table 5.1 Complete list of words used to address gestural antagonism showing stress with IPA diacritic and syllable structure	162
Table 6.1 Summary of constraints on all consonants and the nature of contextual pharyngealization in each	205

Abstract

An essential characteristic of continuous speech is the great articulatory and acoustic variability of sounds, owing to the influence of surrounding segments. This segmental variability, referred to as coarticulation or assimilation, has been a central issue in coarticulation models and is subject to the influence of phonological and physiological constraints. This dissertation examines these constraints using emphasis (known as either pharyngealization or uvularization) as a case study and presents a novel investigation of the C-to-C effect of emphasis, i.e. the effect induced by an emphatic consonant on the preceding consonants. It addresses three core questions. The first is whether the contextual effect exerted by pharyngealized consonants is constrained by phonological factors such as phoneme contrast – i.e. if consonants that are contrastive in pharyngealization (e.g. /t s/) and those that are not (e.g. /n l/) demonstrate similar assimilatory effects. Another central question is whether and to what extent assimilation is influenced by the gestural conflict between the target and trigger consonants that arises from the physiological lingual constraints. This is addressed by comparing the assimilation in consonants varying in the degree of lingual constraints: highly constrained [g ʃ j], moderately constrained [χ] and unconstrained [b f]. The third question is whether the assimilatory effect of pharyngealization is a phonetic or phonological process. Ultrasound and audio data were obtained from 15 adult native speakers of Eastern Peninsular Arabic spoken in certain regions of the Arabian Gulf states as they read target words in a carrier sentence. Tongue contours from different time points of the target and trigger consonants were traced manually.

Articulatory data confirmed that emphatics /tʰ sʰ/ are pharyngealized (not uvularized) by retracting the tongue root towards the back pharyngeal wall and, consequently, depressing and flattening the tongue body and dorsum. The results of polar SS-ANOVA and root retraction ratios

revealed that phoneme contrast does not inhibit assimilation: coronals exhibit categorical effects whether they are contrastive in pharyngealization (e.g. /t s/) or non-contrastive (e.g. /n l/). Assimilation leads to a neutralization of their pharyngealization contrast in /s/ for many speakers and in /t/ for some speakers as the assimilated /t/ and /s/ become articulatorily indistinct from pharyngealized /t^h/ and /s^h/, respectively. For other speakers, distinctiveness is maintained by completely resisting assimilation in /t/. A gestural conflict between pharyngealization and another articulation can result in two possible outcomes depending on its severity. An extreme conflict between pharyngealization and tongue root advancement in [g ʃ j] results in a complete resistance to coarticulation. A moderate conflict between pharyngealization and dorsum raising and retraction in [χ] leads to resistance to pharyngealization in some cases and to a reduction of dorsal raising to allow some degree of coarticulatory root retraction in other cases. Finally, the lack of lingual conflict in [b f] allows the consonants to coarticulate unrestrictedly with the pharyngealized consonant. The realization of the effect as phonetic or phonological is contingent on the presence of either phonological and/or physiological constraints. It is categorical (assimilation or resistance) in the case of phonemic contrast as in /t s/ and gestural antagonism as in [χ g ʃ j], and gradient in the absence of both constraints as in [b f l].

The results support the predictions of Recasens' (1997) degree of articulatory constraints model, suggesting a correlation between the degree of articulatory constraints and the magnitude of coarticulatory pharyngealization. They are, however, only partially consistent with Cohn's (1990) target-interpolation model. One issue with this model is that it does not account for the speaker variability in assimilating some consonants such as [b f l], which are underlyingly unspecified for pharyngealization and unconstrained articulatorily. The model, however, accounts for the role of phonetic constraints in restricting phonetic interpolation, as in [g ʃ j].

CHAPTER 1

Physiology of Speech and Pharyngealization

1.1 Introduction

Normal speech is subject to various constraints imposed by the human speech apparatus and the phonological system of the language being spoken. These constraints should ideally motivate the production of a speech signal that serves its communicative purpose without exerting too much pressure on the speech apparatus or violating the phonological distinctiveness of sounds. Their effect on speech becomes particularly evident when examining coarticulation which is, broadly speaking, the articulatory and acoustic variability of sounds caused by neighboring segments in the speech chain.

Coarticulation can be constrained by inherent physiological constraints on the coarticulated segment ([Gick & Wilson, 2006](#); [Recasens, Pallarès, & Fontdevila, 1997](#); [Recasens & Rodríguez, 2016](#)). These physiological constraints, such as those related to limitations of the mechanical properties of the articulators and antagonism of different muscular activities, give rise to a potential conflict between coarticulation and the articulatory demands of the segment. This, in turn, may hinder the consonant's ability to undergo coarticulation. For instance, assume that a consonant produced with a protracted tongue root such as [j] precedes another consonant which requires the root to be retracted such as the pharyngeal [ħ] (as in the sequence [jaħ]). In that context,

coarticulation induced by [h] would require a retraction of the tongue root in [j], which conflicts with its canonical position. The question that is worth investigating here is: what are the strategies that are employed to resolve these conflicting articulatory demands? This is a core question in this dissertation.

Coarticulation can also be subject to the influence of a language's phonological system ([Manuel, 1999](#); [Manuel & Krakow, 1984](#)). The crucial issue here is that, in order to deliver a clear spoken message, speakers must maintain the distinctiveness between different sounds in the speech chain. At the core of this claim is that languages differ widely on what sounds qualify as distinctive and what sounds do not. One means of preserving this distinctiveness would be to articulate sounds with as little variability as possible – to avoid coarticulating them with other segments so that they maintain their canonical phonetic properties. In practice, coarticulatory variability is the norm in the world's languages, but different languages tolerate different degrees of departure from the canonical forms of specific segments. Therefore, coarticulation is expected to be tied to the language-particular system of distinctiveness, and this is one of the central issues in this dissertation. The question that thus arises is how the systems of contrast of given languages constrain coarticulatory variation.

Segmental effects are classified as either phonetic or phonological according to the magnitude and temporal extent of the effect on the coarticulated segment compared to the coarticulation trigger. Phonetic coarticulatory effects are characterized by their gradient nature whereas phonological assimilatory ones are categorically similar to the assimilation trigger spatially and temporally. This dissertation addresses the effect of physiological and phonological constraints on the phonetic-phonological nature of the effect.

The interaction between coarticulation and those constraints has led to the emergence of various hypotheses and models that were devised to account for the differing coarticulatory patterns observed across the world's languages. Two such models are particularly relevant to this dissertation: one is based on phonological features and proposed by [Cohn \(1990\)](#) and another is based on the articulation requirements of sounds and developed by [Recasens et al. \(1997\)](#).

In this dissertation, Arabic emphasis is used to investigate physiological and phonological constraints on coarticulation. Emphasis refers to a contrastive secondary posterior constriction found in some dental /t d ð/ and alveolar /s/ consonants (called 'emphatics'). In most Arabic varieties, plain /t d ð s/ contrast with emphatic /tˤ dˤ ðˤ sˤ/, whereas other dental and alveolar consonants such as /θ n z l/ do not have emphatic counterparts. There is an ongoing debate among phoneticians about the exact location of this posterior constriction and the key articulator responsible for it. Some laboratory evidence suggested that the dorsum is responsible for emphasis and argued that it is a form of uvularization ([Zawaydeh & de Jong, 2011](#)) whereas other evidence suggested that it is a type of pharynx constriction produced by retracting the tongue root ([E. Al-Tamimi & Heselwood, 2011](#); [Embarki, Ouni, Yeou, Guilleminot, & Al-Maqtari, 2011](#)). The exact lingual configuration during the production of emphasis in Eastern Peninsular Arabic is an important question addressed in this dissertation.

While traditional Arabic grammarians exclusively described the emphasis contrast, modern linguistics has shed light on 'emphasis spread', a coarticulatory effect induced by emphatic consonants. It is a bidirectional process affecting the preceding segments as in /batˤ/ 'ducks' and the following ones as in /sˤab/ 'he poured', where emphasis spreads to the preceding and following segments /a/ and /b/. Previous studies diverge in the exact description of emphasis spread. Most phonological studies assume that the feature responsible for the emphasis contrast, usually

[+RTR], spreads to all segments except those that are specified as [+high] or palatals ([Alwabari, 2018b](#)) and over a specified domain that differs cross-dialectically ([Card, 1983](#); [S. Davis, 1993, 1995](#); [Zawaydeh, 1998](#)). Articulatory studies ([such as Ali & Daniloff, 1972](#)) suggest that emphatic consonants induce a coarticulatory backing gesture (rather than feature) during the production of neighboring vowels. As far as I know, no previous study has looked at the effect of emphasis spreading on consonants.

The reason for choosing emphasis to examine physiological and phonological constraints on coarticulation is that emphasis spread is likely to be constrained both phonologically and physiologically. Phonologically, it is contrastive in subset of coronal consonants while other coronals are not, which allows an assessment of the role of contrast in coarticulation. Physiologically, at least two or three regions of the tongue participate directly in emphasis: the tongue body, dorsum and root. This entails that a large set of consonants can be potentially in articulatory conflict with emphasis. The location of the conflict (among other factors) introduces varying degrees of conflict, each of which can stimulate a different resolution strategy.

Taking all these issues into account, this dissertation addresses the following core questions:

- 1) Articulatorily, is distinctive emphasis pharyngealization or uvularization?
- 2) Is emphasis spread on the preceding consonants phonetic or phonological?
- 3) Is emphasis spread constrained by contrast and how?
- 4) Is it subject to physiological constraints? What are the strategies employed to resolve articulatory conflict between emphasis spread and the articulatory demands of the affected consonants?

The questions are addressed using ultrasound imaging, a technique that allows a visual examination of the shape of the tongue at a high temporal resolution.

Before delving into these theoretical and empirical issues about coarticulation, this chapter presents an anatomical account of the speech apparatus (§1.2) which is essential in describing the articulatory mechanisms involved in the production of emphasis (§1.3). The discussion will then focus on the Arabic dialect that is used in this dissertation (§1.4). The structure of the dissertation is provided in §1.5.

1.2 Anatomy of the Speech Apparatus and Speech Physiology

This section focuses on two articulators that are of particular relevance to the articulation of pharyngealization: the tongue and pharynx. Although this dissertation is primarily concerned with the position rather than the internal anatomical structure of these articulators, understanding their musculature and articulatory movements is important to fully grasp two key concepts: the articulation of pharyngealization and how possible physiological dependence between tongue regions could impact assimilatory pharyngealization.

1.2.1 The tongue. The tongue is entirely composed of muscles and is therefore capable of performing an abundance of complex, highly controlled movements. As a muscular-hydrostat ([Kier & Smith, 1985](#)), its musculature is saturated with aqueous liquids and, therefore, is characterized by an important biomechanical property of hydrostats, volume preservation. This means that a change in one dimension of the tongue results in a compensatory change in at least one other dimension. This concept is essential to my research hypothesis about the possible antagonism between two regions of the tongue, the root and dorsum, as the advancement of the tongue root results in a compensatory increase in the dorsum height because of the volume of the tongue must be preserved.

Muscles in general have two points of attachment: an origin which is fixed and an insertion which moves relatively freely upon contracting the muscle. The muscular architecture of the tongue is relatively complex and consists of two types of muscles differing their origin: extrinsic muscles originating from the surrounding skeletal structures and inserting into the tongue, and intrinsic muscles originating from and inserting into different parts of the tongue ([E. P. Davis, 1999](#); [Gick, Wilson, & Derrick, 2013](#); [Hixon, Weismer, & Hoit, 2014](#); [Kent, 1998](#); [Sanders & Mu, 2013](#); [Seikel, King, & Drumright, 2000](#); [Stone et al., 2018](#); [Zemlin, 1998](#)). Most lingual muscles are paired, which means the tongue can be divided into two lateral halves.

1.2.1.1 Extrinsic muscles. Extrinsic lingual muscles include the genioglossus, styloglossus, palatoglossus and hyoglossus (Figure 1.1), each of which is explained here in terms of its structure and its role in moving the tongue (according to [Gick et al., 2013](#); [Hixon et al., 2014](#); [Kent, 1998](#); [Sanders & Mu, 2013](#); [Seikel et al., 2000](#); [Stone et al., 2018](#); [Zemlin, 1998](#)).

The genioglossus is the largest and strongest extrinsic muscle; it forms the bulk of the tongue. It is a flat and fan-shaped muscle that originates as three bundles of fibers from the mental spine (i.e. a small projection of bone along the midline of the posterior aspect of the mandible): the lower bundle inserts into the tongue root, the middle into the juncture between the dorsum and blade, and the upper into the tongue tip. The genioglossus is responsible for a diverse set of tongue motions according to which bundle of fibers is contracted. Possible movements include: apex protrusion and root advancement so as to draw the tongue anteriorly by contracting the posterior genioglossus fibers (as in the production of high vowels); lowering and retracting the front portion of the tongue by contracting the anterior genioglossus fibers; bunching the tongue body upward by contracting the middle bundle (as in English bunched [ɪ]); and pulling the midline of the tongue

downward to make a longitudinal trough-like depression on the tongue surface (i.e. midline groove) by contracting both anterior and posterior genioglossus fibers.

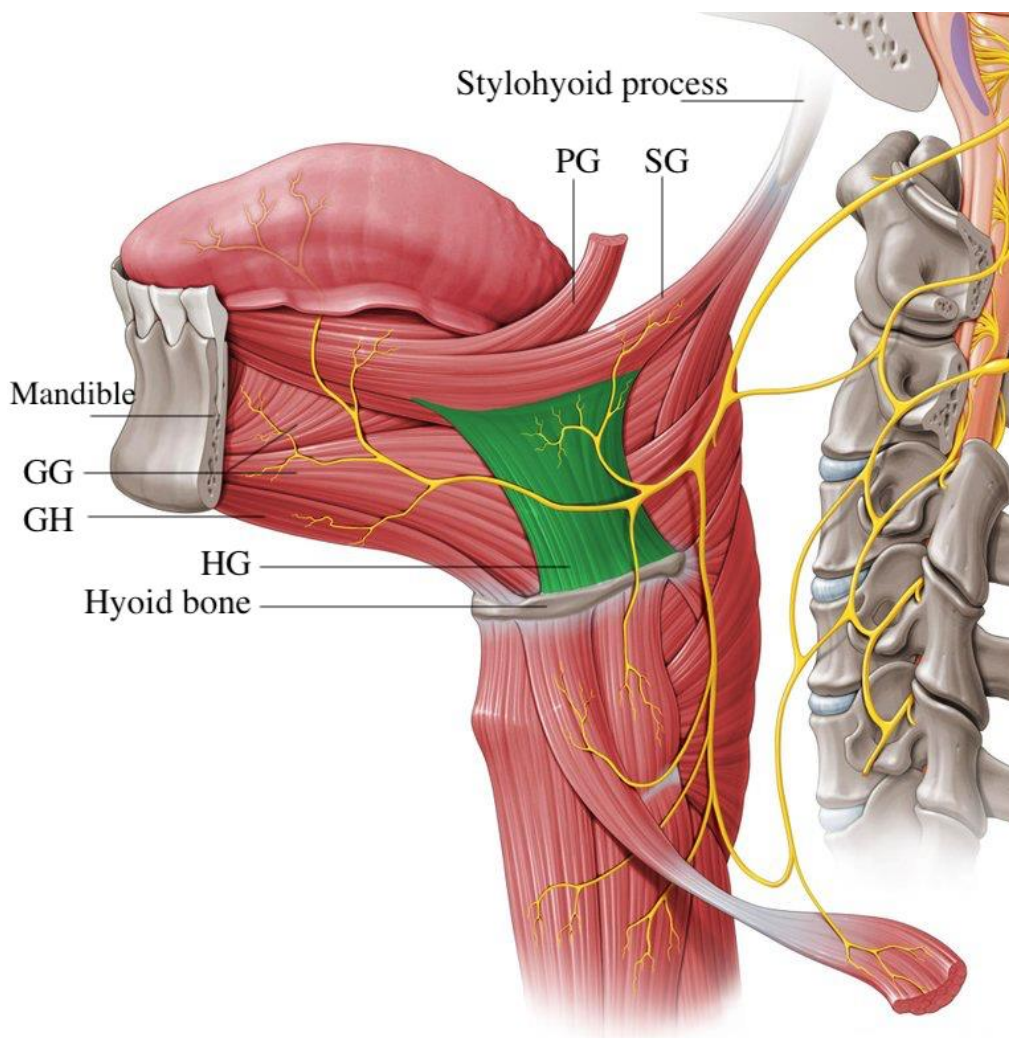


Figure 1.1. The speech apparatus showing the cervical vertebrae and extrinsic lingual muscles: GG= genioglossus, HG= Hyoglossus, PG= palatoglossus and SG= styloglossus. GH= geniohyoid (a non-lingual muscle). Reproduced from [Crumbie, Salvador, and Rad \(2019\)](#).

The styloglossus muscle originates from the styloid process of the temporal bone. Some fibers course anteriorly and downward and insert into the sides of the tongue root where they

radiate toward the midline and the tongue body. Other styloglossus fibers interdigitate with intrinsic muscles and the hyoglossus muscle. Contraction of the styloglossus can have several motor consequences, including pushing the dorsum backward in antagonism with the genioglossus, raising the lateral portions of the tongue, drawing the tip toward the sides of the oral cavity and pushing the dorsum upward as in high back vowels. Note that the latter motor consequence is now controversial according to a study by [Takano and Honda \(2007\)](#) showing MRI evidence that a tendon limits the role of the styloglossus in pulling the tongue upward.

The origin and insertion points of the palatoglossus is a matter of inconsistency among studies. It can be thought of as originating from the lower surface of the soft palate with its fibers spreading downward and forward and inserting laterally into the sides of the tongue ([either into the body as suggested by Gick et al., 2013](#); [or into the root as suggested by Hixon et al., 2014](#)). Alternatively, it can be viewed as originating from the tongue lateral edges and inserting into the soft palate, in which case it is called glossopalatine. The exact point of attachment of the palatoglossus in the soft palate varies across individuals: it is either near the hard palate or near the uvula, which could affect its role in velar versus lingual movements ([Kuehn & Azzam, 1978](#)). The palatoglossus can pull the tongue mass backward and, by a simultaneous contraction of the two sides of palatoglossus, it can make a longitudinal groove on the tongue surface.

The hyoglossus forms a thin, quadrilateral sheet of muscle that originates from the greater cornua of the hyoid bone (i.e. the greater cornua is horn-like pair of bones at the back of the hyoid bone and project upward). Its fibers course vertically and insert into the sides of the posterior half of the tongue. The hyoglossus muscle is responsible for retracting and depressing the tongue, pulling the sides downward (in direct antagonism with palatoglossus) as well as elevating the hyoid bone.

1.2.1.2 Intrinsic muscles. Intrinsic lingual muscles, which form much of the tongue mass, are responsible for squeezing and deforming the tongue into different shapes, but not for pulling the tongue toward hard skeletal structures ([Gick et al., 2013](#)). Thus, these muscles provide the tongue with the biomechanical properties of muscular hydrostats. Biomechanically, deforming an intrinsic muscle requires a second muscle either to stiffen a portion of the tongue, as in bending, or to provide muscular support, as in trilling the tongue tip ([Stone et al., 2018](#)). There are four intrinsic muscles making up two pairs: a pair of longitudinal muscles and another of interwoven transverse and vertical muscles (Figure 1.2).

There are two longitudinal muscles: superior and inferior. The superior longitudinal muscle is a flat, broad and thin layer of fibers beneath the tongue surface. Its fibers originate from the root and extend anteriorly to the tip. Because fibers near the midline of the tongue are thicker and because some fibers at the tip run obliquely toward the lateral boundaries of the tongue, the superior longitudinal muscle is capable of curling the sides of the tongue (or the lateral regions) upward. As with any muscular hydrostatic organ, contracting this longitudinal muscle causes the front of the tongue to bend as in retroflex sounds. It can also shorten the tongue and give it a concave shape. An important characteristic of the superior longitudinal muscle is that, because its structure consists of short overlapping fibers, it can provide functional control of local portions of the tongue independently. For example, retroflexion is achieved by a specific functional segment (i.e. front part of the superior longitudinal muscle) rather than the whole muscle (Stone, Epstein & Iskarous, 2004). This minimizes articulatory conflict and, thus, resistance to assimilation in speech production.

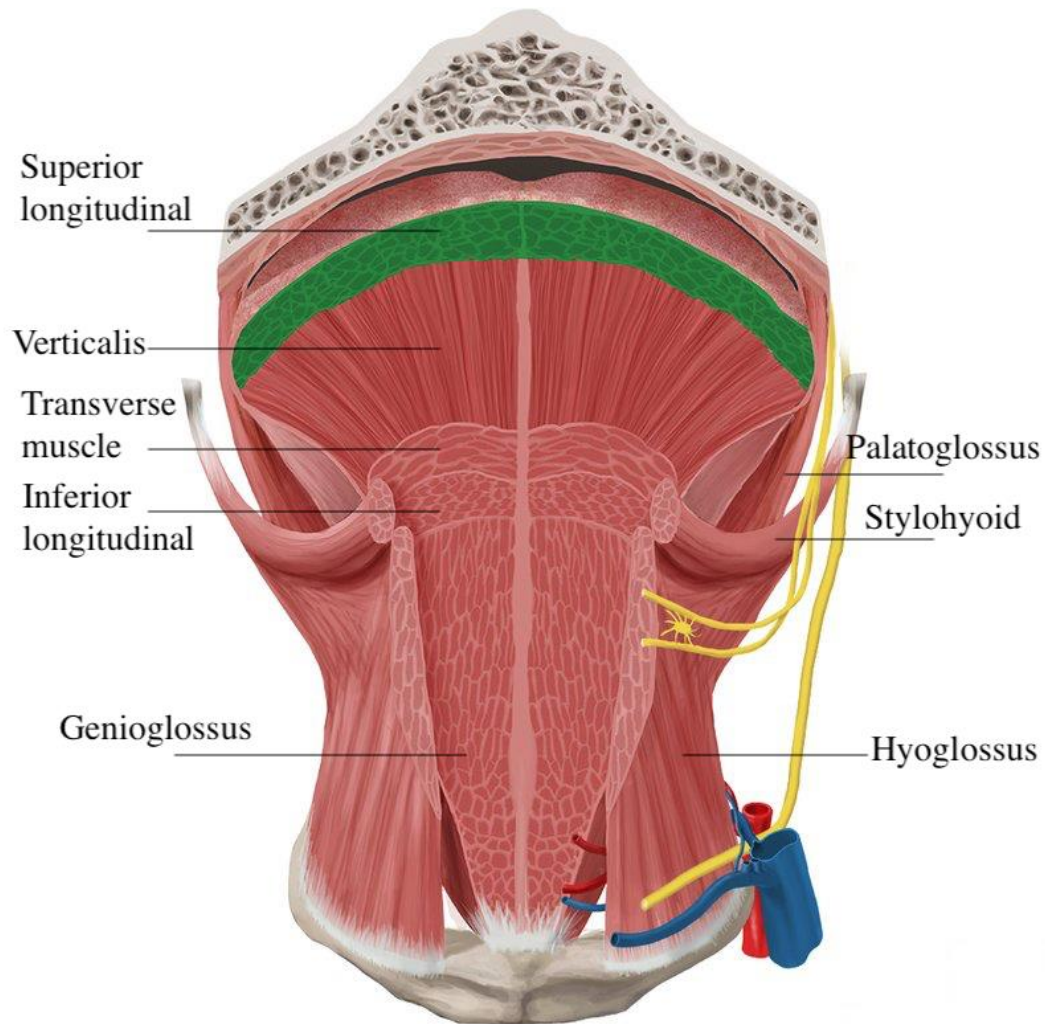


Figure 1.2. A coronal view of the tongue illustrating the intrinsic and some extrinsic lingual muscles.

Reproduced from [Crumbie et al. \(2019\)](#).

The inferior longitudinal muscle, which lies near the undersurface of the tongue on both sides of the genioglossus, originates from the hyoid bone near the root and courses forward to insert into the lower surface of the tongue tip. Some biomechanical functions of the inferior longitudinal muscle are antagonistic to that of the superior longitudinal muscle. Namely,

contracting the anterior portion of the inferior longitudinal muscle results in curling the tongue tip downward, and contracting the entire inferior longitudinal muscle gives the tongue a convex shape.

The transversus, which forms layers from the root to the tip, radiates horizontally and laterally from the midline of the tongue to either side. The transversus squeezes the tongue from side-to-side and, thus, protracts and thickens the tongue mass.

The verticalis, as the name suggests, forms layers of fibers that interdigitate with transversus fibers and runs vertically upward between the inner surfaces of the superior and inferior longitudinal muscles. The verticalis is narrower at the top than the bottom and contracting it elongates, widens and flattens the tongue.

1.2.2 The pharynx. The pharynx, an approximately 12-cm tube of tendons and muscles extending from the base of the skull to the cricoid cartilage anteriorly and the sixth cervical vertebra posteriorly, comprises three cavities: the nasopharynx, oropharynx and laryngopharynx. The nasopharynx is situated behind the nasal cavity and above the velum and, since the velum is mobile, the lower boundary of the nasopharynx is fluctuating. The laryngopharynx, the lowest division of the pharynx, is surrounded by the hyoid bone superiorly, the base of the cricoid cartilage inferiorly and the back surface of the tongue root and the epiglottis anteriorly. Compared to the static nasopharynx, the oropharynx, which is situated between the nasopharynx at the top and the laryngopharynx at the bottom, is relatively dynamic due to the mobility of the surrounding tongue base and velum.

The pharyngeal musculature comprises three main muscles, each consisting of a set of individual muscles (Figure 1.3). First, the superior pharyngeal constrictor muscle, the uppermost muscle of the pharynx that encircles the upper pharynx, consists of four individual muscles. These muscles form a cluster of left-and-right bundles originating from different points at the front of the

pharyngeal tube and radiating backward to insert into the posterior midline of the pharynx. The middle pharyngeal constrictor muscle which consists of two individual muscles originates from the hyoid bone and the stylohyoid ligament and courses backward toward the posterior midline of the pharyngeal tube. Finally, the inferior pharyngeal constrictor muscle, which comprises two individual muscles, arises from the thyroid and cricoid cartilages. Contracting any of these three pharyngeal constrictor muscles pulls the posterior pharyngeal wall forward and the lateral pharyngeal wall inward, thus reducing the cross-sectional dimension of the pharyngeal lumen. Therefore, they participate in constricting the pharynx horizontally but this is independent of the retraction of the tongue root, which also constricts the pharynx. There is disagreement in studies on pharyngealization about whether the pharynx is constricted via the retraction of the tongue root solely or also with the aid of the posterior and lateral walls.

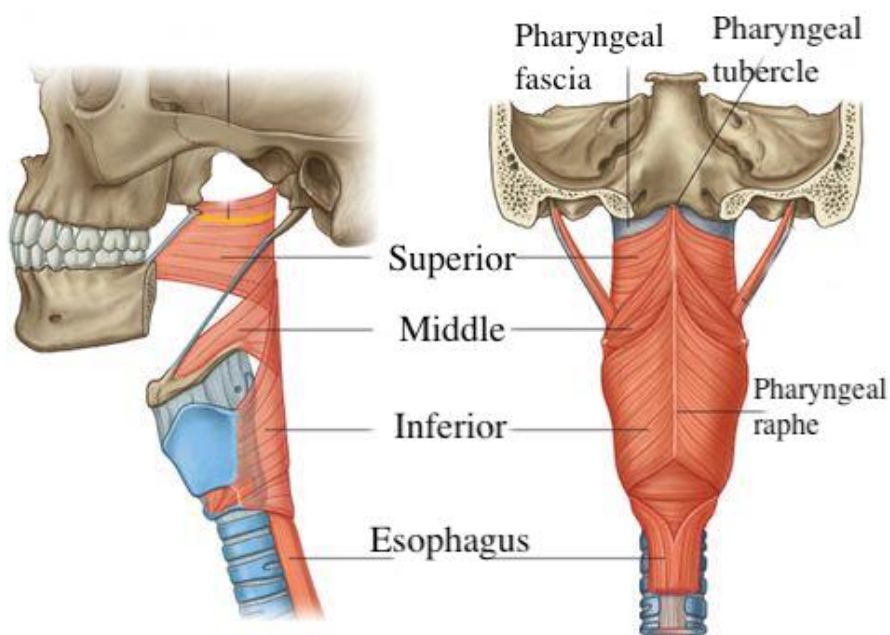


Figure 1.3. Sagittal (left) and posterior (right) view of the speech apparatus showing the superior, middle and inferior pharyngeal constrictors.

1.3 Emphasis

1.3.1 Articulation of emphasis. Arabic emphatic consonants involve a primary constriction in the anterior region of the oral cavity, mostly with the tongue tip and blade against the alveolar ridge or the dental-alveolar area, simultaneously with a post-velar constriction (the emphasis proper). There is, however, a lack of consensus among researchers about: 1) the configuration of the posterior region of the vocal tract during emphasis and more specifically about the precise location of the secondary constriction, 2) its resemblance to pharyngeal and uvular consonant articulation, 3) the larynx position during emphasis, and 4) the role of the dorsum, tongue root, epiglottis and hyoid bone in emphasis. This has sparked off an intense debate about whether emphasis is realized as uvularization or pharyngealization. The following discussion will provide a complete picture of all articulatory gestures involved in emphasis and distinguish it from the production of uvular and pharyngeal consonants after describing their production. To facilitate the comparison, Table 1.1 below presents a summary of the articulation of uvulars, pharyngeals and emphatics.

1.3.1.1 Articulation of uvulars and pharyngeals. A number of parameters are involved in the articulation of Arabic pharyngeal consonants /ħ ʕ/: the degree of pharyngeal stricture, its location and the role of the epiglottis and other parts of the speech apparatus.

In terms of degree of constriction, the voiceless pharyngeal /ħ/ is consistently described as a fricative. Its voiced counterpart /ʕ/ seems more variable: while it is usually described as a fricative, the pharyngeal diameter in its production is greater than in /ħ/ in endoscopic images of Jordanian-Arabic speakers ([Zawaydeh, 2003](#)), and it has thus been described as an approximant in this dialect ([Esling, 1999](#); [Laufer & Condax, 1981](#)). In contrast, it is described as a stop in the Iraqi dialect ([Al-Ani, 1970](#); [Butcher & Ahmed, 1987](#)).

Apart from the differing degrees of stricture, both /ħ, ʕ/ are produced at the same location with the same mechanism. Their constriction location is reported to be at the laryngo-pharynx which makes them the most posterior sounds in Arabic ([Delattre, 1972](#); [Ghazeli, 1977](#); [Laufer & Baer, 1988](#)). This constriction is formed at the level of the fourth cervical vertebra as can be seen in cinefluorographic images of a Tunisian speaker ([Ghazeli, 1977](#)) or at the level between the second and third cervical vertebrae as videofluoroscopic data of Jordanian speakers suggest ([Heselwood & Al-Tamimi, 2011](#)). In any case, the narrowest constriction in pharyngeals is consistently between the back pharyngeal wall and the epiglottis, which is tilted downward and backward with the aid of the suprahyoid muscle ([as illustrated in Figure 1.4, reproduced from Lawson, Stuart-Smith, Scobbie, & Nakai, 2018](#)). The horizontal compression of the pharyngeal cavity is achieved by the retraction of the epiglottis and the tongue root ([Ali & Daniloff, 1972](#); [Laufer & Condax, 1978, 1981](#); [Shar & Ingram, 2011](#)). Some studies revealed that the tongue root movement is accompanied with an inward movement of the posterior pharyngeal wall ([Ghazeli, 1977](#)) or of the lateral pharyngeal walls moving the faucal pillars close to each other ([Catford, 1977](#)). Many studies reported that the epiglottal movement is accompanied with a retraction of the tongue root ([Altairi, Brown, Watson, & Gick, 2017](#); [Catford, 1977](#); [Esling, 1999](#)) whereas others argued that the epiglottis moves independently of the tongue root due to the contraction of the aryepiglottic muscle and, thus, must be regarded as the primary articulator in pharyngeals ([Heselwood & Al-Tamimi, 2011](#); [Laufer & Baer, 1988](#); [Laufer & Condax, 1978, 1981](#)). Given this role of the epiglottis in the production of emphatic consonants, it has been proposed that they should be classified in the IPA as epiglottal ([Esling, 1999](#); [Laufer & Baer, 1988](#)) or epiglottopharyngeal ([Catford, 1977](#)).



Figure 1.4. Vocal tract configuration in producing pharyngeals /ħ ʕ/ showing pharyngeal compression.

Reproduced from [Lawson et al. \(2018\)](#).

Other articulatory gestures associated with the production of pharyngeals include: dorsum lowering ([Altairi et al., 2017](#)); and arytenoids raising ([Esling, 1999](#)); and larynx raising ([Catford, 1977](#); [Esling, 1996](#); [Ghazeli, 1977](#); [Heselwood, 2007](#); [Laufer & Baer, 1988](#)). Larynx raising (estimated as 7mm compared to its position in other speech sounds) is dependent on tongue root retraction according to [Esling \(2005\)](#). Therefore, larynx raising and tongue root retraction are assumed to be associated, to various degrees, not only with pharyngeals, but also with other consonants produced with a constricted pharynx (such as emphatics). Less agreed-upon is the posture of the aryepiglottic folds which can be either: a) completely constricted between the arytenoids and the base of the epiglottis ([Esling, 1996](#)); b) kept open along their length as in the voiced approximant /ʕ/ or constricted with a slight aperture between the arytenoids as in the

voiceless fricative /ħ/ ([Esling, 1999](#)); or c) trilling ([Hassan, Esling, Moisik, & Crevier-Buchman, 2011](#)).

The articulation of uvulars is less complex and more agreed-upon than the articulation of pharyngeals (e.g. [Bin-Muqbil, 2006](#); [Delattre, 1972](#); [Ghazeli, 1977](#); [Laufer & Baer, 1988](#)). There are three uvular consonants in Arabic dialects: the voiceless stop /q/ (sometimes realized as velar [g]) and the two fricatives /ʁ χ/. The latter fricative is also realized in some dialects as a voiceless velar fricative [x].

All uvular allophones are articulated with the dorsum retracted backward towards the upper pharynx, as illustrated in Figure 1.5 reproduced from [Lawson et al. \(2018\)](#). According to [Delattre \(1972\)](#), the formation of the uvular constriction involves a semi-circular movement of the dorsum which first retracts horizontally towards the back pharyngeal wall and then slides vertically and upward along the pharynx. Uvulars vary in the degree of their dorsal constriction with the stop /q/ having the smallest aperture. The uvular /ʁ/ involves trilling achieved by pushing the uvula downward against the raised dorsum so that the pulmonic airflow, under Bernoulli effect, causes it to oscillate. It should be noted here that in the articulation of /ʁ/ the uvula is not necessarily curled as shown in Figure 1.5, but is rather slightly moved downward according to [Delattre \(1972\)](#) and [Ghazeli \(1977\)](#).



Figure 1.5. Vocal tract configuration in producing uvular /ʁ/ (left) and /q/ (right) showing uvular aperture and no epiglottal retraction. Reproduced from [Lawson et al. \(2018\)](#).

1.3.1.2 Articulation of *emphatics*

1.3.1.2.1 Emphasis as uvularization. There has long been a debate as to whether emphatics are the result of uvularization or pharyngealization. The belief that emphatics are a form of uvularization (or sometimes velarization) has originated from the work of the famous 8th-century Arabic grammarian, Sibawayh, and was supported by more recent linguistic research. These studies are based on either formal phonological analysis ([McCarthy, 1994](#); [Shahin, 1996, 1997](#)), acoustic or articulatory empirical evidence ([Al-Ani, 1970](#); [Altairi et al., 2017](#); [Bin-Muqbil, 2006](#); [Ghazeli, 1977](#); [Herzallah, 1990](#); [Zawaydeh, 1997, 1998, 1999](#); [Zawaydeh & de Jong, 2011](#)), and the focus in this section will be on articulation-based studies.

Two parameters have informed the uvularization account of emphatics: the location of the posterior constriction along the vertical dimension of the vocal tract and the overall configuration of the vocal tract compared to pharyngeals and uvulars. For example, the cinefluorographic studies of [Ghazeli \(1977\)](#) and [Zawaydeh \(1999\)](#) revealed that the posterior articulation of emphatics is

formed at the level of the second cervical vertebra just like that of uvulars, whereas it is at the fourth for pharyngeals. This secondary constriction is made by a backward retraction of the tongue dorsum against the uvula accompanied by a depression of the palatine dorsum, a configuration quite similar to uvulars ([as shown in Figure 1.6, reproduced from Bin-Muqbil, 2006](#)). Although the pharyngeal tube is compressed in uvularization compared to the plain consonants in Figure 1.6, this compression seems to be the by-product of tilting the dorsum upward and backward and it is less tight than in the pharyngealization account. The retraction of the tongue dorsum against the uvula is the greatest for the stop /tʕ/ (10mm displacement compared to its plain counterpart /t/) and the least for the fricative /sʕ/ (6mm displacement compared to plain /s/). Two other arguments against the characterization of emphatics as pharyngealized include the lower pharyngeal constriction and the role of the larynx. Unlike pharyngeals, no horizontal contraction of the laryngo-pharynx was reported in the production of emphatics, either by retracting the tongue root or by advancing the posterior pharyngeal walls ([Ghazeli, 1977](#)). That is, the width of the lower pharynx in emphatics /tʕ sʕ ʔʕ/ is similar to that of their plain counterparts /t s ʔ/. Furthermore, although pharynx contraction is physiologically associated with larynx raising ([Esling, 2005](#)), uvularization does not involve any vertical displacement of the larynx or the hyoid bone. Also, no lip protrusion was observed in [Ghazeli \(1977\)](#) whereas [Lehn \(1963\)](#) reported labialization associated with these sounds. The vocal tract configuration reported in [Ghazeli \(1977\)](#) was found to be consistent with Iraqi emphatics by [Ali and Daniloff \(1972\)](#). In brief, the formation of a secondary articulation at the uvula in emphatics and the lack of vertical and horizontal contractions of the laryngo-pharynx (either by retracting the tongue root and raising the larynx) makes emphatics comparable to uvulars and distinct from pharyngeals.

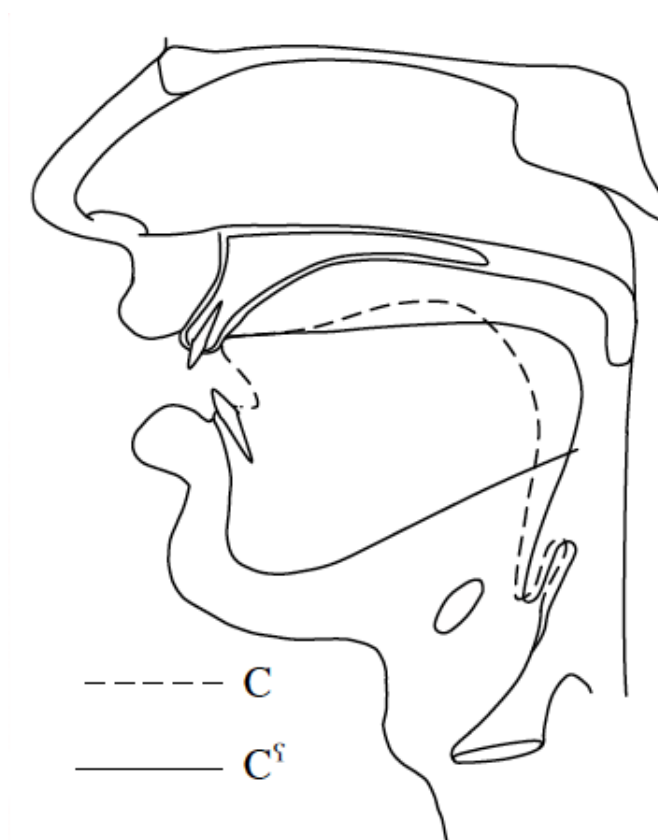


Figure 1.6. Vocal tract configuration in producing emphatics as uvularized showing the secondary articulation at the uvula. Reproduced from [Bin-Muqbil \(2006\)](#).

1.3.1.2.2 Emphasis as pharyngealization. A recent line of research views emphasis as pharyngealization based on findings of various articulatory imaging studies conducted on different dialects. The most salient parameters for characterizing it as pharyngealization rather than uvularization are the involvement of the tongue root and epiglottis and the resulting laryngo-pharyngeal compression and laryngeal raising. Although many studies of emphasis agree on the pharyngealization account of emphasis, there is some inconsistency on the precise posture of some articulators.

In general, emphatics are produced with a configuration qualitatively similar to pharyngeals (with some exceptions) but with a less extreme displacement of the various parts of

the speech apparatus. The key parts that previous research has examined include: tongue dorsum and root, uvula, pharynx, epiglottis, lingual muscles, hyoid bone and various laryngeal structures.

Findings converge on the role of the tongue root, which retracts towards the posterior pharyngeal wall, and of the epiglottis, which approximates the arytenoid cartilages, thus constricting the pharynx horizontally ([J. Al-Tamimi, 2017](#); [Ali & Daniloff, 1972](#); [Catford, 1977](#); [Hassan & Esling, 2007, 2011](#); [Laufer & Baer, 1988](#); [Laufer & Condax, 1981](#); [Shar & Ingram, 2011](#); [Zeroual & Clements, 2015](#); [Zeroual, Esling, & Hoole, 2011](#)). The tongue root, therefore, is deemed a main articulator in producing the secondary articulation of pharyngealization ([Ali & Daniloff, 1972](#)). The cross-sectional compression along the pharyngeal tube may be achieved by retracting the tongue root only ([Ali & Daniloff, 1972](#)) or accompanied by an inward movement of the posterior and lateral pharyngeal walls ([E. Al-Tamimi & Heselwood, 2011](#)) via the pharyngeal constrictor muscles. The role of the lateral and posterior walls is a matter of dispute, not only in the production of pharyngealized consonants but also in pharyngeal consonants. This pharyngeal compression is attained with a series of successive gestures as described in [E. Al-Tamimi and Heselwood \(2011\)](#) and [Hixon et al. \(2014\)](#). When the tongue root is pushed backward against the anterior surface of the epiglottis, it covers the epiglottic valleculae (i.e. lateral channels running between the epiglottis and the tongue root) and, consequently, forces the epiglottis to retract backward. The upward-pointing epiglottis deflects backward against the posterior pharyngeal wall and, then, downward, making a seal over the laryngeal aditus (an action similar to swallowing whereby the epiglottis protects the pulmonary airway). The tongue root is allowed to retract against the epiglottis, thus, reducing the pharyngeal volume by the deactivation of the genioglossus and geniohyoid muscles, whose contraction advances the tongue and expands the pharyngeal tube in plain consonants ([Kuriyagawa, Sawashima, Niimi, & H., 1988](#)).

There is far less agreement about the posture of the tongue dorsum in emphasis. It was described as: 1) raised up in ultrasound images in multiple dialects including Yemeni, Palestinian, Egyptian and Moroccan ([Zeroual et al., 2011](#)); 2) depressed in the cinefluorography and videofluoroscopy of the Iraqi and Libyan dialects ([Ali & Daniloff, 1972](#); [Lardi, 1983](#)); or 3) retracted in EMA trackings of Tunisian ([Embarki, Ouni, Yeou, et al., 2011](#)). The dorsal backing observed in the latter study by Embarki and colleagues starts during the preceding vowel and persists to the following vowel and is more robust in Modern Standard Arabic – as spoken by native speakers of Yemeni, Kuwaiti, Jordanian and Moroccan – than in the regional dialects.

Even when they agree that emphasis is a form of pharyngealization, studies do not necessarily obtain consistent results about the exact position of the dorsum and the lingual muscles at play. That is, the dorsal elevation found in some studies ([Altairi et al., 2017](#); [Bin-Muqbil, 2006](#)) is attributed to the contraction of the styloglossus which pulls the tongue body upward and backward ([Hixon et al., 2014, p. 228](#)). On the other hand, the dorsal depression in emphatics and pharyngeals found in other studies (e.g. [Ali & Daniloff, 1972](#); [Esling, 2005](#); [Lardi, 1983](#)) has been explained in relation to the hyoglossus, whose contraction results in lowering the mass of the tongue and retracting it towards the posterior pharyngeal wall. Although both the styloglossus and hyoglossus cause retraction of the tongue, their antagonistic effect on tongue dorsum elevation may have contributed to the observed inconsistency of the dorsum height in emphatics.

The larynx is raised in emphatics, just as it is in pharyngeals ([F. Al-Tamimi, Alzoubi, & Tarawnah, 2009](#); [F. Al-Tamimi & Heselwood, 2011](#); [Hassan & Esling, 2007](#); [Zeroual & Clements, 2015](#)). The only exception to this is [Shar and Ingram \(2011\)](#) which did not find any laryngeal displacement and [Hassan and Esling \(2011\)](#) which reported laryngeal lowering in emphatics. This laryngeal elevation is reportedly accompanied with an elevation of the hyoid bone ([F. Al-Tamimi](#)

[et al., 2009](#)) and, therefore, both participate in a two-dimensional compression of the pharyngeal tube together with the retracted tongue root. As in pharyngeals, the results on the status of the aryepiglottic folds in emphatics are not conclusive as they were found either constricted ([Hassan & Esling, 2007](#)) or not ([Zeroual et al., 2011](#)).

Overall, although pharyngealized consonants are produced with qualitatively similar mechanisms to pharyngeal consonants, there are still some quantitative differences between them. For example, larynx raising in pharyngeal consonants is greater than in emphatics – i.e., by 7 mm in [Ghazeli \(1977\)](#) and 4-6 mm in [E. Al-Tamimi and Heselwood \(2011\)](#). Furthermore, root retraction and the subsequent cross-sectional pharyngeal compression are quantitatively different from that observed in pharyngeal consonants in terms of the degree, location and variability of the constriction. Firstly, the pharynx is more cross-sectionally constricted in pharyngeal consonants than in emphatics by retracting the tongue root and epiglottis to a greater extent ([Laufer & Baer, 1988](#)). Moreover, the point of the tightest constriction in pharyngeal consonants is at the fourth cervical vertebra ([Ghazeli, 1977](#)) whereas it is at the second or third cervical vertebrae in emphatics ([E. Al-Tamimi et al., 2009](#); [E. Al-Tamimi & Heselwood, 2011](#)). Also, according to [Laufer and Baer \(1988\)](#), unlike pharyngeal, which show a consistent degree of pharyngeal constriction in all contexts, the degree of constriction in emphatics is affected by the consonant's manner of articulation (with stop /tʰ/ having tighter constriction than fricatives /sʰ ʃʰ/) and the context vowel (with consonants preceding /a/ having narrower constriction than in /i/ context).

Table 1.1

Comparison of the articulation of uvulars, pharyngeals and emphatics

	Uvulars	Pharyngeals	Emphatics /tˤ sˤ dˤ ðˤ/	
	[q χ ʁ]	[ħ ʕ]	Uvularization	Pharyngealization
Lips	Not tested	Not tested	Labialization in some [21] but not [12] all studies, especially /dˤ/	Labialization in some but not all studies, especially /dˤ/
Tongue body	Depressed, flattened	Depressed, flattened	Depressed, flattened	Depressed & flattened [1, 18] or retracted [8]
Dorsum	Tilted backward and upward against uvula [7]	Tilted backward and downward due to root gesture [2]	Tilted backward and upward against uvula [5]	Tilted backward [8]
Tongue root	No movement	Substantial retraction [12, 19]	May and may not retract very slightly [12]	Moderate retraction [1, 14, 20, 24]
Uvula	Varying degrees of uvular constriction by dorsal raising	No role	Uvular constriction. Greater for /tˤ/ than /sˤ/ [5]	No role

	Uvulars	Pharyngeals	Emphatics /tʕ sʕ dʕ ðʕ/	
	[q χ ʁ]	[ħ ʕ]	Uvularization	Pharyngealization
	and/or uvular lowering [7, 12]			
Pharynx	Slight compression of oropharynx	Extreme compression at oropharynx, narrowest at epiglottis due to retracted root & advanced posterior wall [6, 12]	Moderate compression of oropharynx	Moderate compression of oro- and laryngo-pharynx, narrowest at epiglottis [3]
Epiglottis	No role	Retracted (in)dependently from tongue root posture [17, 19, 20]	No role	Retracted [20]
Narrowest constriction	At uvula	At epiglottis, at the 4 th cervical vertebra [12]	At uvula or oro-pharynx, 2 nd cervical vertebra [12, 22]	Between 2 nd & 3 rd cervical vertebra [3, 4]
Hyoid bone	No role	Raised	No role	Raised [3]
Larynx	No displacement	Raised by 7mm, compressing pharynx vertically [9, 11, 12, 16]	No displacement	Raised by 4-6mm, compressing pharynx [4]

	Uvulars	Pharyngeals	Emphatics /tʕ sʕ dʕ ðʕ/	
	[q ɣ ʁ]	[ħ ʕ]	Uvularization	Pharyngealization
Aryepiglottic folds	Not tested	Constricted [9], open [10] or trilling [15]	Not tested	Constricted [13] or open [24]

[1] [Ali and Daniloff \(1972\)](#); [2] [Altairi et al. \(2017\)](#); [3] [E. Al-Tamimi and Heselwood \(2011\)](#); [4] [E. Al-Tamimi et al. \(2009\)](#)

[5] [Bin-Muqbil \(2006\)](#); [6] [Catford \(1977\)](#); [7] [Delattre \(1972\)](#); [8] [Embarki, Ouni, Yeou, et al. \(2011\)](#)

[9] [Esling \(1996\)](#); [10] [Esling \(1999\)](#); [11] [Esling \(2005\)](#); [12] [Ghazeli \(1977\)](#)

[13] [Hassan and Esling \(2007\)](#); [14] [Hassan and Esling \(2011\)](#); [15] [Hassan et al. \(2011\)](#); [16] [Heselwood \(2007\)](#)

[17] [Heselwood and Al-Tamimi \(2011\)](#); [18] [Lardi \(1983\)](#); [19] [Laufer and Condax \(1978\)](#); [20] [Laufer and Baer \(1988\)](#)

[21] [Lehn \(1963\)](#); [22] [Zawaydeh \(1999\)](#); [23] [Zeroual et al. \(2011\)](#); [24] [Zeroual and Clements \(2015\)](#)

1.3.1.2.3 Variation in emphasis articulation. The discussion in the previous two sections has highlighted two different mechanisms of producing emphasis. This inconsistency may have stemmed from methodological issues or dialect differences. It can be an artifact of the experimental procedures and equipment used in investigating emphasis articulation. For instance, articulation instruments that cannot capture the movement of the posterior portion of the tongue, such as EMA, may lead to assuming that the dorsum is the principal articulator and may conclude that the more posterior portion of the tongue does not play a role in emphasis articulation. On the other hand, other instruments that track the movement of both the anterior and posterior regions, such as MRI, can make more accurate conclusions about the movement of the tongue root. Such articulation inconsistency can be, according to [Khattab, Al-Tamimi, and Heselwood \(2006\)](#), due to the dialect under investigation, the phonological context, speakers' gender, and sociolinguistic factors. To address both possibilities, methodological and dialectal bases, a survey of some existing studies that found either uvularization or pharyngealization is presented in Table 1.2

Table 1.2

Survey of studies supporting uvularization and pharyngealization

Studies supporting uvularization			
Study	Dialects & speakers	Method(s)	Major findings
Al-Ani (1970)	MSA spoken in Iraq	X-rays	Constriction between the root and upper pharynx at the 2 nd cervical vertebra
Herzallah (1990)	Palestinian	X-rays	The dorsum approximates the upper pharynx (similar to uvulars). So, emphasis is both velarization & pharyngealization
Bin-Muqbil (2006)	MSA (5 Central Saudi speakers)	Acoustic	Dorsum retraction similar to uvulars
Ghazeli (1977)	One Southern Tunisian	Cinefluorography, nasal airflow	Retraction of the back of the tongue to upper pharynx (at 2 nd cervical vertebra), depression of the palatine dorsum, oral expansion & pharyngeal narrowing
Zawaydeh (1999)	Ammani-Jordanian	Endoscopy ¹	Retracting the back of the tongue towards the upper pharynx
Altairi et al. (2017)	3 Yemeni 2 Egyptian 2 Saudi 1 Palestinian	Ultrasound	Dorsum raising, root retraction & body lowering similar to uvulars
Studies supporting pharyngealization			
Ali and Daniloff (1972)	3 Iraqi	Cinefluorography	Root retraction, dorsum depression, oral expansion, no role of posterior pharyngeal wall or hyoid bone

E. Al-Tamimi et al. (2009)	4 Jordanian	Videofluoroscopy ²	Root retraction into oropharynx (at 2 nd cervical vertebra), raising hyoid bone and larynx. No role of velum or uvula
E. Al-Tamimi and Heselwood (2011)	9 Jordanian	Nasoendoscopy ³ , videofluoroscopy	Root retraction, constriction at the epiglottis, inward movement of pharyngeal walls, larynx raising & dorsum lowering
Hassan and Esling (2007, 2011)	2 Iraqi	Laryngoscopy ⁴	Root retraction, larynx raising & aryepiglottic sphinctering
Laufer and Baer (1988)	2 Palestinian 1 Lebanese 1 Iraqi	Cineradiography	Root and epiglottis retraction, laryngopharyngeal constriction
Zeroual et al. (2011)	2 Moroccan	EMA ⁵ , endoscopy, ultrasound	Retracting the epiglottis and tongue at the oropharynx

¹ Endoscopy consist in visualizing/filming the movement of the pharynx with a light and camera inserted through the oral cavity.

² Videofluoroscopy consist in making X-ray movies of the larynx and supra-laryngeal tract along the sagittal and/or coronal plane.

³ Nasoendoscopy consists in inserting a flexible tube with light at the end through a nostril and positioning it at the top of the pharynx. It collects transverse-view images of the pharynx with a low sampling rate.

⁴ Laryngoscopy consists in inserting a fiberoptic tube in the pharynx through a nostril. A camera attached to the fiberoptic scope captures transverse images of the larynx.

⁵ EMA stands for electromagnetic articulography, a tool for studying speech kinematics. It involves placing sensors on the tongue, lips and jaws to track their movement with a high temporal resolution as they move in an electromagnetic field.

It appears from the survey that neither of these potential factors can account for the inconsistency in producing emphasis. No method yields the same findings across different studies: for example, the fluoroscopy which can image the entire vocal tract demonstrated that emphasis is articulated by retracting the dorsum ([Herzallah, 1990](#)) whereas the same method used in [F. Al-Tamimi et al. \(2009\)](#) showed that it is the tongue root that is retracted. Moreover, the most studied dialects, Jordanian and Palestinian, are found to have emphasis as both uvularization ([Altairi et al., 2017](#); [Zawaydeh, 1999](#)) and pharyngealization ([F. Al-Tamimi et al., 2009](#); [Laufer & Baer, 1988](#)).

What this survey reveals is that emphasis can be realized as either (or *both*) uvularization and pharyngealization. Finding evidence of either articulation in a given dialect does not mean that all speakers of this dialect must articulate emphasis in the same way. Therefore, no cross-dialectal and cross-methodological consensus can be reached among studies about the articulation of emphatics. Rather, there seems to be variation in the realization of emphatics and this variation does not appear to be dialectically-based.

1.3.2 Acoustics of emphasis. Acoustic correlates of emphasis are based on the examination of vowels preceding or following emphatics using traditional formant frequency measures and Bark differences between formants ([J. Al-Tamimi, 2017](#); [Hassan & Esling, 2007, 2011](#)), locus equation ([Embarki, Ouni, & Salam, 2011](#); [Embarki, Ouni, Yeou, et al., 2011](#); [Embarki, Yeou, Guilleminot, & Al-Maqtari, 2011](#); [Yeou, 1997](#)) and spectral tilt ([J. Al-Tamimi, 2015](#)).

Acoustically, the coarticulation of emphasis is reportedly not limited to the vowel immediately adjacent to the emphasis source, but can extend all the way to the word boundary or beyond ([Bukshaisha, 1985](#)). This has been found in regional Arabic dialects other than the Arabian Peninsula dialects, which have not been examined extensively ([S. Davis, 1993, 1995](#); [Younes, 1994](#)). In the current study, the distance between assimilated consonant and the emphasis source

is strictly limited to one inter-consonantal vowel. In addition, acoustic evidence suggests a difference between regressive and progressive coarticulation on vowels within the same dialect in terms of their temporal span. For example, [Jongman, Herd, Al-Masri, Sereno, and Combest \(2011\)](#) found that regressive coarticulation in Jordanian Arabic extends throughout the entire vowel (i.e. at vowel onset, mid-point and offset) whereas progressive coarticulation was limited to the vowel onset and mid-point. Since the focus of the present study is not on these directional asymmetries, the assimilatory effects addressed here are limited to the regressive coarticulation only. The reason for selecting regressive rather than progressive coarticulation is that the former has been shown to be more dominant than the latter in some dialects, both in terms of its temporal domain and the magnitude of its effect on neighboring vowels ([Jongman et al., 2011](#); [Younes, 1994](#)).

Two major themes in the acoustic literature are relevant to this dissertation. The first is the association between the acoustic properties of coarticulatorily emphatic vowels and the articulation of emphasis, and more specifically the acoustic evidence supporting either the uvularization or the pharyngealization account of emphatics. The other theme is related to general characteristics of coarticulation – e.g. temporal extent – which are relevant to the experimental design in this dissertation. Both themes are discussed below.

1.3.2.1 Acoustic evidence for pharyngealized or uvularized constrictions. The formant frequencies of vowels can be taken as indices of the vocal tract shape and of the precise location of a secondary articulation. Acoustic evidence could thus help us determine if emphatics are uvularized or pharyngealized.

The most robust acoustic correlate of emphasis across dialects is a F2 drop in the vowels surrounding an emphatic consonant. This drop, which suggests the backing of pharyngealized vowels, has been consistently reported in various dialects including Palestinian ([Card, 1983](#)),

Egyptian ([Kahn, 1975](#)), Moroccan ([J. Al-Tamimi, 2017](#); [Yeou, 1997](#); [Zeroual et al., 2011](#)) Tunisian ([Ghazeli, 1977](#)), and Jordanian ([E. Al-Tamimi & Heselwood, 2011](#); [Jaber, Omari, & Al-Jarrah, In press](#); [Jongman et al., 2011](#); [Khattab et al., 2006](#); [Zawaydeh, 1997, 1998, 1999, 2003](#); [Zawaydeh & de Jong, 2011](#)). F2 lowering is associated with a constriction in the laryngopharynx, and is consistent with the videofluoroscopic findings of [E. Al-Tamimi and Heselwood \(2011\)](#), which showed a pharyngeal compression at the same area. An F2 drop is also observed in vowels adjacent to uvulars and pharyngeals; however, the effect of emphatic consonants is much greater than the effect of uvulars ([Zawaydeh, 1998](#); [Zawaydeh & de Jong, 2011](#)). This can be due to emphatic consonants having a more extreme and tighter pharyngeal constriction than uvulars. Also, as [Zawaydeh \(1998\)](#) suggested, emphatic consonants tend to induce a great effect on neighboring vowels to distinguish them from the vowels preceding their non-emphatic counterparts.

Emphasis also raises F1 on neighboring vowels, which may be an effect of the rounding and protrusion of the lips. This is consistent with the articulatory evidence of [Zeroual et al. \(2011\)](#) which reported labialization in producing emphatic consonants. Although not as strong as the effect on F2 lowering, the increase in F1 has been reported in the majority of the acoustic studies.

1.3.2.2 Other factors affecting the properties of coarticulated vowels. Various other factors have been shown to affect the magnitude of emphasis spread, such as the quality and length of vowels, the manner of articulation of the emphatic consonant, and the temporal extent of coarticulation. It is crucial to mention them here as they influenced the stimuli selection criteria used in this dissertation. Since these effects have only been measured acoustically so far, I have chosen to discuss them in this section.

Different vowel qualities exhibit different degrees of sensitivity to emphasis ([Card, 1983](#); [Embarki, Ouni, Yeou, et al., 2011](#); [Ghazeli, 1977](#); [Jongman et al., 2011](#)). Anticipatory and carry-over coarticulation are more evident on /æ/ which incurs greater F2 drop and F1 increase than on /i u/, which seem more resistant to coarticulation. Also, in regressive coarticulation, emphasis effects were observed on F1 and F2 of /æ/ regardless of its distance from the pharyngeal consonant, whereas the acoustic effect rapidly diminished on /i u/ as distance from the emphatic consonant increased.

Vowel length also affects emphasis spreading as its effect on F1 and F2 is stronger at the midpoint of short than long vowels ([Jongman et al., 2011](#)), owing to the fact that the temporal interval between the emphatic consonant and the vowel mid-point is smaller in short vowels. This is true at vowel onset, midpoint and offset and for both directions of coarticulation.

The manner of articulation of emphatic consonants induces asymmetrical coarticulatory effects on vowel formants (F2 in particular); namely, stops /t^s d^s/ influence the neighboring vowels more significantly than fricatives /ð^s s^s/ ([Jongman et al., 2011](#)). In fact, the words with emphatic fricatives in [Jongman et al. \(2011\)](#) also included the pharyngeal consonant /ħ/, which induces coarticulatory effects similar to (and even more extreme than) the effects of the emphatic consonants ([Hassan & Esling, 2007](#)). Words with the emphatic stops /t^s d^s/ were not subject to this confounding effect from the pharyngeal /ħ/. This means that the extreme coarticulation that [Jongman et al. \(2011\)](#) observed in these words may not be attributed to the manner of articulating the emphatic consonant; rather, they can be caused by an additive coarticulatory effect of the pharyngeal consonant.

1.4 Dialect Typology

Emphasis in this dissertation is tested with speakers of Eastern Peninsular Arabic. This loosely-defined variety is spoken in several Arabian Gulf states along the Eastern coast of the Arabian Peninsula: Bahrain, Kuwait, Qatar, United Arab Emirates and the cities of Dammam, Ahsa, Qatif, which are part of Saudi Arabia. It does not include varieties spoken in Oman, and excludes all the dialects spoken in other parts of the Arabian Peninsula (Yemen and the Central and Western Regions of Saudi Arabia).

In fact, the Arabian Peninsula as a whole is relatively a large area with 24,354,444 nationals (who speak the regional dialects) and therefore it is bound to dialectical diversity. Previous studies diverge on what counts linguistically as a dialect in the Arabian Peninsula with some considering all varieties spoken in the area as one dialect and combining divergent varieties such as Hijazi and Bahraini as one dialect ([Ingham, 1982](#)) and others associating every small region in one of these countries with a different dialect, thus, ignoring the commonalities between these varieties and others ([Al-Mozainy, 1981](#); [Alammar, 2017](#); [Alhammad, 2014](#); [Alzaaq, 2017](#); [Holes, 1983](#)). In this dissertation, I take an intermediate stance between these two extreme classifications of dialects and argue that some of those sub-varieties along the east coast of the Arabian Peninsula converge on some isoglosses. My classification is quite similar to that of [Holes \(1990\)](#) with the exception that I included the varieties spoken in the cities of Qatif and Dammam (i.e. the former is identical to that spoken in Bahrain and similar to other Eastern Peninsular varieties, and the latter is very similar to the varieties in Ahsa and Qatif as many speakers immigrated from these two cities).

Several typological commonalities characterize the varieties of Eastern Peninsular Arabic and distinguish it from the other dialects spoken in other parts of the Arabian Peninsula. Taking into consideration the consonant phoneme inventory presented in Table 1.3, several allophonic

variations of these consonants are shared among the sub-varieties comprising the Eastern Peninsular Arabic.

Table 1.3

Chart of the Arabic consonant inventory, adapted from [Thelwall and Sa'adeddin \(1999\)](#)

	Bilab.	Lab. Dent	Dent.	Alve.	Post- alveolar	Pala.	Velar	Uvul.	Pharyn.	Glott.
Plosive	b		t d t ^ʕ d ^ʕ				k	q		ʔ
Nasal	m		n							
Trill				r						
Fricative		f	θ ð s z ʃ θ ^ʕ s ^ʕ				x ɣ		ħ ʕ	h
Affricate					ḏʒ					
Approx.						j	w			
Lateral				l						

Unlike other varieties spoken in the Gulf states, Eastern Peninsular varieties often palatalize the velar stop /k/ into an affricate [tʃ] before front vowels (i.e. k ~ tʃ / _i, such as [ki:s]-[tʃi:s]’bag’) and to either the fricative [ʃ] or the affricate [-tʃ] in the 2nd person feminine suffix e.g. (i.e. -ek ~ -eʃ ‘you-FEM’, such as [sʰənek]-[sʰənetʃ] or [sʰəneʃ] ‘your plate’).

Although the uvular stop /q/ is realized as [g] in all Peninsular dialects, Eastern Peninsular Arabic differs from the rest by another allophonic variant, [ḏʒ], in contiguity with front high

vowels e.g. [qiddæm]-[d̥iddæm] ‘in front’ (Holes, 1990, pp. 262). The variant [d̥] is not used in other Peninsular varieties such as Hijazi, Omani and Yemeni (with the latter two avoiding both variants of /q/: [g] and [d̥]). Certain sociolinguistic classes of speakers tend to stick to the allophonic variant [g] and avoid [d̥].

Furthermore, the phoneme /d̥/ (not the allophonic variant of /q/) is realized as [j] in all phonetic contexts in Eastern Peninsular varieties. Words such as [d̥ʒa:r]-[ja:r] ‘neighbor’ are evident example of this alternation (i.e. d̥ ~ j) which is avoided by speakers of some social classes. This is not attested in other Peninsular varieties.

In addition to these consonants, the Eastern Peninsular dialect differs from other Peninsular dialects in the allophonic forms of some vowels in certain contexts. For example, the vowel /a/ verb-initially and in the 1st person singular pronoun /anʒ/ ‘I’ surfaces as [a] (thus, yielding [anʒ]).

1.5 Structure of the Dissertation

This dissertation was designed to investigate the interaction between assimilation (emphasis spread in particular), on the one hand, and the constraints imposed by the tongue physiology and by the phonology on the other hand. It also aims to examine the phonetic versus phonological characterization of emphasis spread. All of these goals are addressed in an ultrasound experiment and discussed in three separate chapters.

CHAPTER 2 offers an in-depth discussion of coarticulation in three respects so that I make reliable predictions about emphasis spread based on existing evidence from the literature. It first defines the two segmental processes, assimilation and coarticulation, that are the core of this dissertation, and establishes certain criteria for distinguishing them. The chapter then reviews hypotheses and experimental evidence about two types of constraints on assimilation: phonemic contrast and articulatory constraints. This is followed by a discussion of coarticulation models with

a focus on two recent models built around these hypotheses: Cohn's target-interpolation model and Recasens' Degree of Articulatory Constraints model. The chapter concludes with a set of predictions about emphasis spread in Arabic deriving from these models.

CHAPTER 3 addresses the first two research questions posited earlier: 1) whether distinctive emphatics are pharyngealized or uvularized, and 2) whether emphasis spread is phonetic or phonological. The chapter presents the methodology used for collecting and analyzing the ultrasound data used in Chapter 3-5. The results and discussion sections in Chapter 3 present and interpret the findings on the lingual configuration used to produce contrastive emphasis with a special focus on the tongue root, which is found to be the principal articulator. It also presents the results on the extent to which nearby consonants were affected by emphasis.

CHAPTER 4 addresses the third question of this dissertation: whether (and how) emphasis spread is influenced by the contrast system of Arabic by comparing two sets of coronals: those that contrast in /t s/ and those that do not /n l/. Then, the results which show that contrast does not restrain assimilation but rather enhances categorical assimilation on various consonants are presented.

CHAPTER 5 follows a similar structure as the previous chapter and investigates the counteraction of physiologically-based gestural conflict and assimilation. It compares assimilation in three sets of consonants varying in the degree of lingual constraints and gestural conflict with emphasis: labials [b f], velars [x g] and palatals [j ʃ]. The results of the lingual configuration on each of these three sets is presented according to the degree of gestural antagonism which is associated with different conflict resolution strategies.

CHAPTER 6 summarizes the findings of the preceding chapters. It argues for the interaction between phonological and physiological constraints, particularly in cases where both were

predicted to induce contradictory effects on assimilation. In addition, the chapter compares the present findings on emphasis with the predictions derived from the two models of coarticulation presented in Chapter 2.

CHAPTER 2

Coarticulation: Theories and Predictions

2.1 Introduction

An essential characteristic of continuous speech is the great articulatory and acoustic variability of sounds, owing to the influence of surrounding segments. This segmental variability, referred to as coarticulation or assimilation, is characterized by its susceptibility to various factors – including, but not limited to, phonological and physiological constraints – and has motivated the emergence of recent coarticulation models. The aim of these models, which are discussed in this chapter, is to predict the characteristics of coarticulation and assimilation and establish an association between the invariant units of representation and variable phonetic forms. The models make completely different predictions about what segments can be assimilated with an emphatic trigger given the fact they vary in their unit of representation – articulatory gestures or segmental features; therefore, a segment that is featurally compatible with emphasis may not be gesturally compatible.

This chapter offers an empirically grounded discussion of two types of factors that can constrain assimilation, phonological contrast and articulatory conflict (§2.3), as well as a review of two types of coarticulation models, featural and gestural (§2.4). The chapter concludes with detailed predictions about emphasis spread in Arabic based on the two types of constraints and the two coarticulation models. However, before reviewing this literature in detail, it is important to

define coarticulation and distinguish it from the other type of contextual variability, assimilation (§2.2).

2.2 Assimilation as Feature Spreading and Coarticulation as Coproduction

A traditional assumption in the literature is that there is a binary divide between two assimilatory processes: assimilation proper, a phonological and categorical process implemented by feature spreading, and coarticulation, a phonetic and gradient process that derives from the coproduction of adjacent segments¹.

It is believed that assimilation entails feature spreading/copying whereby an assimilated phoneme acquires a new feature from a trigger phoneme by spreading ([Clements, 1985](#)), leading to a substantial change of the property of the assimilated segment throughout its entire duration ([Keating, 1990](#)). Feature spreading is traditionally associated with cognitive processing and mental representation of speech sounds whereas coproduction is related to the neuromotor execution of speech sounds. Feature spreading, as proposed by [Daniloff and Hammarberg \(1973\)](#) and [Hammarberg \(1976\)](#), is assumed to take place before sending neural commands for the articulation of the target sound to the speech mechanism, resulting in a mental reorganization of the articulatory planning of the target. The rationale for attributing feature spreading to the brain is twofold. First, it involves uncovering the upcoming (distant) trigger and anticipating its structure and articulation – a process that requires access to memory and not only to the subcortical and peripheral systems responsible for the execution of articulator movements. (Early models assumed that only phonological processing took place in the cortex while motor control was more peripheral, occurring in the subcortical system. However recent models assume that motor output is the result of the interaction between the sub-cortex and certain lobes of the cortex). Second, the fact that the

¹ For a survey of gradient and categorical assimilatory processes, see [Cohn \(1993\)](#) and [Flemming \(2001\)](#).

spreading occurs multiple segments away from the trigger phoneme ([Daniloff & Moll, 1968](#); [Moll & Daniloff, 1971](#)), which is more than what is physiologically required, suggests that feature spreading is a cognitive process.

It is this featural adjustment of the canonical form of segments (indicative of cognitive preplanning) that differentiates feature spreading from coproduction, as proposed by [Fowler \(1980, 1985\)](#). In coproduction, the canonical forms of segments are not altered but their intrinsic temporal extent allows them to overlap in time, i.e. phonemes are not modified by the context but rather coproduced, resulting in an overlapping production ([Bell-Berti & Harris, 1977, 1981, 1982](#); [Bell-Berti, Krakow, Gelfer, & Boyce, 1995](#); Bell-Berti & Krakow, 1991; [Solé, 1992, 1995](#))². This leads to a fine-grained and gradient variation of a given phonetic property that lasts only for a portion of the segment and cannot be captured by discrete and timeless phonological representations. Coproduction would thus take place at the peripheral, neuromuscular system ([Farnetani & Recasens, 1999](#)). However, [Wood \(1996\)](#) argues that coproduction (and thus coarticulation) could be cognitive, just like feature spreading, in that the temporal extension of a gesture from the trigger may be preplanned rather than mechanical.

To illustrate the levels of processing at which feature spreading and coproduction could take place, consider the example of Swedish retroflexion, presented in [Wood \(1996\)](#). In some Swedish dialects, when the retroflex /r/ is followed by one or more alveolar consonants (e.g. ‘Tors stad’), these consonants are retroflexed and the retroflex phoneme /r/ is elided (e.g. [tu:ʂ ʂtɑ:d] where the three consecutive alveolars are retroflexed and the vowel blocks the effect from extending into the final [d]). If this is feature spreading, it should involve transferring the retroflexion from the /r/ to the following alveolar consonants before motor production. In a classic

² These studies, which led to the formulation of the ‘time-locked’ model of coarticulation, were based on the view of coarticulation as coproduction and they considered timing to be intrinsic to the motor organization of segments.

coproduction account, the retroflex gesture of /r/ would be treated as intrinsically long. It would overlap with the alveolar gestures of /s/ and /t/. However, an alternate account in which coproduction is preplanned may more adequately explain why the alveolar gesture does not spread in all contexts and why it is blocked by the vowel. This type of extrinsic co-production would involve scanning the upcoming phonemes for their (in)compatibility with retroflexion and rephasing the temporal extent of retroflexion accordingly, a process that must be taking place at the cognitive level.

In fact, there have recently been proposals for a gradient phonology that assume that phonological patterns and processes are characterized with continuous, rather than discrete, representations ([Cohn, 2006](#); [Scobbie & Stuart-Smith, 2008](#)); however, this dissertation will stick to the traditional association between categoricalness and phonological processes in defining the effect of emphatics on nearby segments because it is still dominant in the field and well-established. We will see in the discussion of the results that it is certainly an oversimplification.

Taking this distinction between coarticulation and assimilation into account, the question that arises is whether the contextual effect induced by emphatics is coarticulation (involving coproduction) or assimilation (a form of feature spreading). In fact, the majority of studies on Arabic emphasis analyze it as a form of feature spreading and, therefore, it is commonly known as ‘emphasis spread’ ([S. Davis, 1995](#); [Halle, Vaux, & Wolfe, 2000](#); [Herzallah, 1990](#); [McCarthy, 1991, 1994](#)). However, this consensus may reflect the phonological focus adopted by most of these studies rather than an empirical reality. Therefore, articulatory investigation of emphasis spread is needed to examine whether it is featural or caused by gestural coarticulation. The contextual effect of emphatics on neighboring consonants can be deemed as assimilatory feature spreading provided

that the target consonant is categorically pharyngealized to the same degree as the trigger to comply with the binary nature of features.

2.3 Factors Affecting Coarticulation

Various factors have been reported to influence coarticulation in one way or another. Prosody is the first one. We know that prosodically strong positions (such as stressed syllables and phrase edges) not only enhance the articulatory prominence of a segment but also increase its coarticulation resistance. Stressed vowels, for instance, are less coarticulated with the neighboring segments and exert greater coarticulatory effects than unstressed vowels ([Beddor, Harnsberger, & Lindemann, 2002](#); [Cho, 2000, 2004](#); [Conklin & Dmitrieva, 2017](#); [de Jong, Beckman, & Edwards, 1993](#); [Fowler, 1981](#); [Magen, 1997](#); [Recasens, 2015](#)). Given this, the stress of both target and source segments in the present study are strictly controlled to avoid any confounds. The two other factors, phonological contrast and physiological constraints, are central to the present study and are discussed in more detail in the following subsections.

2.3.1 Phonological contrast. The effect of phonological contrast on contextual variability in general and coarticulation in particular can be traced back to the 1960s ([Öhman, 1966](#)) but it was later developed by Manuel in the 1980s and 1990s. Manuel's hypothesis about the effect of contrast is based on evidence from comparative studies of V-to-V coarticulation. It was subsequently applied to several coarticulatory phenomena including nasalization cross-linguistically and palatalization in Russian.

2.3.1.1 Manuel's hypothesis about phonological contrast. In order to better understand this hypothesis, which stems from cross-linguistic differences in the distribution of vowels in the vowel space, it is important to highlight a few facts about languages' vowel inventories. Based on a survey of over 300 languages, [Maddieson \(1984\)](#) reported that there is a cross-linguistic

preference for certain vowels and certain patterns of distribution in the acoustic vowel space. For example, vowel inventory sizes vary from three to fifteen vowels in the 317 languages surveyed in [Maddieson \(1984\)](#); however, two-thirds of these have between five and seven distinctive vowels ([Schwartz, Boë, Vallée, & Abry, 1997](#)). Five-vowel inventories tend to consist of peripheral vowels /i e a o u/, with the frequent addition of /ɛ ɔ/ in seven-vowel inventories. The distribution of these vowels is determined not only by the size of inventories, but also by language-specific rules ([Bradlow, 1995](#)). That is, vowels are typically distributed in the vowel space in a way that ensures a sufficient perceptual contrast, according to the dispersion theory ([Liljencrants & Lindblom, 1972](#); [Lindblom, 1986](#)). This claim about vowel dispersion is particularly relevant to Manuel's 'output constraints' hypothesis (also referred to here as the 'contrast hypothesis') on the effect of contrast on coarticulation.

The hypothesis postulates that a language-particular system of "phonetic contrasts" imposes output constraints on the articulation of phones and that these constraints can limit coarticulation ([Manuel, 1987, 1990, 1999](#)). The hypothesis considers every phoneme to be associated with a target acoustic area (rather than specific points) which can accommodate a range of acceptable pronunciations due to contextual variability such as coarticulation. This entails that the magnitude of coarticulation is dependent on the size of the target area of the phoneme affected by coarticulation. Universal principles such as the drive to maintain distinctiveness between phonemes can constrain the size of these target areas and, consequently, the magnitude of coarticulation ([Manuel & Krakow, 1984](#)). Typically, languages with a smaller vowel inventory can allocate more space per vowel than languages with a larger inventory, but the crucial factor affecting coarticulation, according to Manuel's hypothesis, is not the number of vowels *per se*, but rather their distribution in the vowel space and the area allocated to each vowel. Evidence about

the effect of vowel distribution, and not inventory size, on coarticulation is reported in [Manuel \(1987, 1990\)](#), which compared V-to-V coarticulation in three Southern Bantu languages, Shona, Ndebele and Sotho. The first two languages have a five-vowel system /i e ɑ o u/, but have a similar vowel distribution in the high-back region as Sotho, which has seven vowels /i e ε ɑ ɔ o u/; therefore, the three languages are not expected to show any difference in coarticulating high-back vowels. However, the low- and mid-vowel region is less crowded in Shona and Ndebele (in which it is occupied by only the low-central /ɑ/) than in Sotho where the same vowel /ɑ/ is in a close proximity to /ε ɔ/. According to Manuel, due to this distributional difference and the smaller space allocated to /ɑ/ in Sotho, this language exhibits considerably smaller anticipatory coarticulation than the other two languages with a larger acoustic space for the same vowel. The minimal coarticulatory variability observed in Sotho (which is limited to /ɑ/ raising along the central region and not backward or forward raising towards /ε/ or /ɔ/, respectively) functionally preserves phonemic contrast and prevents acoustical and perceptual confusion of the vowel with other vowels that are close to it. A similar pattern is observed in Swahili and Shona, where vowels are more widely dispersed and, therefore, undergo greater coarticulation than English, which has a more crowded vowel space that allows little variability within the small vowel regions ([Manuel & Krakow, 1984](#)). Likewise, V-to-V coarticulation is more restricted in English VCV sequences than in Japanese, which has a sparser vowel system consisting of only /i e ɑ o u/ ([Magen, 1984](#)).

2.3.1.2 Limitations of the contrast hypothesis. One of the limitations to the hypothesis is that contrast fails to predict some variability and coarticulatory phenomena. The hypothesis cannot account for the neutralization of contrast resulting from place assimilation in many languages, for example, ‘can play’ [kæmpleɪ] where the alveolar nasal /n/ is realized as bilabial [m]. In addition, although Korean has a relatively more crowded vowel space, consisting of eight vowels, compared

to the five-vowel system of Japanese, no coarticulatory difference between the two languages is observed in [Han \(2007\)](#). Moreover, the two five-vowel languages in [Manuel \(1990\)](#) pattern differently, as Ndebele exhibits more coarticulatory vowel fronting than Shona. Phonemic contrast and vowel distribution cannot explain the difference between these two languages which have similar vowel distributions. This is not a violation of the hypothesis, but it suggests rather that phonemic contrast sets specific limits on phonetic variability and languages can coarticulate more or less within these limits. The absence of featural contrast of a given phoneme allows it to coarticulate freely, but it does not necessarily mean that the phoneme must coarticulate. In fact, contrast is one of many factors which all affect the coarticulatory variability of phonemes differently. For example, the difference between Shona and English reported in [Beddor et al. \(2002\)](#) is attributed to phonetic factors other than contrast. Their acoustic findings revealed more extensive carry-over and anticipatory coarticulation in English than Shona, which is at odds with the contrast hypothesis and with the findings of [Manuel and Krakow \(1984\)](#) on the same languages. The difference in the magnitude of coarticulation is attributable to another factor, the longer closure duration of the intervocalic stop in Shona VCV sequences, that restricts the magnitude of coarticulation in this language.

Another limitation is that the series of studies with which Manuel developed the hypothesis are not clearly consistent in defining contrast, as some refer to phonetic differences between surface allophones while others refer to phonemic contrast between underlying phonemes. This seemingly is not an issue detrimental to the validity of the hypothesis because the languages that were addressed have a simple mapping between phonemic and allophonic vowels. Nevertheless, the hypothesis can make contradictory predictions when it comes to cases of complex mapping, such as Cantonese and Beijing Mandarin, depending on whether it considers phonemic or

allophonic vowel inventories. That is, Mandarin has a sparser five-vowel space than Cantonese, which has eight vowels, and the distribution difference is particularly evident at the low vowel region. Both languages, however, have a similar allophonic vowel inventory density. Given this, if phonetic contrast is what imposes the output constraints, the magnitude of V-to-V coarticulation should be similar in both languages. On the other hand, if the phonemic contrast is the key factor constraining coarticulation, coarticulation in Cantonese should be more restricted than Mandarin especially for the low vowel /a/. The acoustic inquiry of [Mok \(2012\)](#) and [Mok and Hawkins \(2004\)](#) shows that phonemic density cannot determine the magnitude of coarticulation because vowels in the sparse phonemic inventory (Mandarin) do not allow more coarticulation than vowels in the dense inventory (Cantonese).

2.3.1.3 Applying the contrast hypothesis to other coarticulation phenomena. The contrast hypothesis was developed based on cross-language comparative studies of V-to-V coarticulation where the vowel inventory size was taken as an index of contrast. Another line of research emerged later and extended the hypothesis to within-language coarticulatory differences, i.e. coarticulation of different phonemes of a given language that are either contrastive or non-contrastive for a given feature. Palatalization and nasalization are two examples of coarticulatory phenomena that illustrate this line of research and are relevant to the present study.

Russian palatalized trills present a complex case of phonemic contrast, as they involve both manner contrast with other coronals (i.e. trills vs. plosives /t tʲ/, fricatives /s sʲ z zʲ ʃ ʃʲ/, nasals /n nʲ/ and laterals /l lʲ/) and palatalization contrast (i.e. plain /r/ and palatalized trills /rʲ/). Most Russian coronals are contrastive for palatalization /tʲ dʲ sʲ ʃʲ zʲ nʲ lʲ rʲ/, the only exceptions being /ʈsʲ ʈʃʲ ʒʲ jʲ/. [Iskarous and Kavitskaya \(2010\)](#) examined the effect of this complex contrastive system on the trills' phonetic variability induced by coarticulation and prosodic boundary

strengthening/weakening. Their acoustic analysis of minimal pairs containing /r/ vs. /r̥/ in different vowel contexts and different positions in the word demonstrated that the coarticulatory variability induced by the vowels [a u i ɪ] on the adjacent trills /r r̥/ (e.g. /rada/ ‘close’ vs. /r̥adom/ ‘close’) is not inhibited by the density of the phonemic contrast. Coarticulation does not neutralize the palatalization contrast. Furthermore, phonemic contrast does not restrain the prosodic variability of trills: there is neither a weakening of the palatalization contrast domain-medially nor a strengthening at the boundaries of the prosodic domain. Taken together, these facts suggest that phonemic contrast does not restrict the phonetic variability caused by a prosodic boundary or coarticulation.

The role of contrast in restricting coarticulatory nasalization varies from one language to another, yet a survey of the existing research demonstrated that the nasalization is independent of contrast in many languages being documented. That is, vowels in languages with no distinctive vowel nasality can be less sensitive to coarticulatory nasalization than languages with distinctive vowel nasality. On the one hand, despite their lack of phonemic nasality, Italian ([Farnetani, 1986, 1990](#)) and Spanish ([Solé, 1992, 1995](#)) exhibit minimal or no contextual nasalization on vowels preceding a nasal consonant. The onset of vowel nasalization (as measured by the velopharyngeal port opening) in these two languages starts late in the vowel, thus rendering the vowels oral for most of their production. On the other hand, languages with contrastive vowel nasalization such as Brazilian Portuguese ([Barbosa & Albano, 2004](#); [Barlaz, Shosted, Fu, & Sutton, 2018](#)); French ([Delvaux, Demolin, Harmegnies, & Soquet, 2008](#); [Desmeules-Trudel & Brunelle, 2018](#); [Dow, 2014](#)) and Lakota ([Scarborough, Zellou, Mirzayan, & Rood, 2015](#)) exhibited varying degrees of coarticulatory nasalization. For example, in Lakota, coarticulation led to a neutralization of phonemic contrast between the high back vowels /i u/ and /ĩ ũ/. Phonetically speaking, this can be

attributable to the claim posited by [Maeda \(1993\)](#) that high back vowels in particular can be nasalized with even a smaller velic opening than that required to nasalize low vowels. This, in turn, leaves less acoustic and articulatory space to produce the oral quality in high vowels.

2.3.2 Physiological constraints. Articulatory constraints on coarticulation originate from competing articulatory demands – in either the manner of articulation, aerodynamic requirements or articulatory gestures – of the contextual and the target segments. The effect of articulatory gestures, which is central to this dissertation, is discussed in more detail in §2.3.2.1.

As for the constraints originating from the manner of articulation, [Recasens and Espinosa \(2009\)](#) compared consonants, such as Catalan /s/ vs. /l/, with the same place of articulation but with different manners of articulation in terms of their sensitivity to vowel-induced coarticulation. Their EPG data revealed that lingual constraints are stronger in fricatives (as in /s/) than laterals (as in /l/). The lingual demands associated with the manner of articulation of the fricative /s/, mainly the formation of a narrow central groove to allow the airflow, is what renders the fricative more constrained and, thus, less variable than the lateral /l/.

As for the aerodynamic constraints, two examples illustrate their impact in restricting coarticulation: nasalization of Sundanese obstruents and palatalization of Russian trills. Nasalization of oral obstruents presents a case where the aerodynamic requirements of the primary consonant can interfere with coarticulation. The velum elevates during the production of oral obstruents in order to allow an accumulation of air pressure sufficient for frication or a stop burst ([Ohala, 1971, 1975](#)). In some languages, such as Sundanese, this aerodynamic constraint causes these consonants to resist nasalization, which requires lowering the velum, as this lowering will allow air to leak through the nasal passage and, thus, reduce the oral air pressure ([Ohala, 1990](#)). The Sundanese words [ɲãhõkɾɲ] “to inform” and [mĩʔãsih] “to love” illustrate this concept: vowels

and pharyngeal and glottal obstruents, which are not subject to this aerodynamic conflict, are progressively nasalized, whereas the oral ones [k] and [s] resist and block nasalization.

Palatalization and coarticulation of palatalized trills present another case of aerodynamic constraints restricting contextual variability and, more specifically, coarticulation. Driven by aerodynamic forces rather than muscular control, trilling requires the vibration of the tongue tip. The tongue tip assumes a position against the alveolar ridge or palate by a muscular contraction which decreases the airflow through the aperture and, consequently, increases the oropharyngeal air pressure behind the linguopalatal constriction. According to the Bernoulli principle, the pressure build-up forces the tip away from the alveolar ridge/palate, thus allowing the pulmonic air out. Subsequent reduction in air pressure forces the tip to bounce back against the palate where the pressure starts to increase. Any deviation in airflow or pressure could be detrimental to this self-sustained oscillation ([Kavitskaya, 1997](#); [Shevelov, 1979](#); [Solé, 2002](#); Spajić, Ladefoged, & Bhaskararao, 1996). The aerodynamic complexity of trills nevertheless has contradictory implications across languages – both phonologically and phonetically. This aerodynamic (and articulatory) complexity makes the Catalan trill /r/ more resistant to vowel-dependent coarticulation, compared to the tap /ɾ/ which is less complex aerodynamically and articulatorily ([Recasens & Pallarès, 1999](#)). In contrast, trilling complexity does not restrict vowel coarticulatory effects on Russian plain and palatalized trills /r rʲ/ which were expected to be highly resistant to coarticulation ([Iskarous & Kavitskaya, 2010](#)). These conflicting behaviours highlight the fact that not only aerodynamic constraints are at play here but that gestural conflict (and phonological contrasts) should also be considered as potential sources of constraints on coarticulation in the vicinity of trills.

2.3.2.1 Defining gestural conflict. Gestural conflict, which is one type of articulatory constraint, occurs when the segments involved in coarticulation require one or more competing articulatory targets (gestures) to be achieved with the same articulator or within the same articulatory zone. For example, the conflict between the dorsal raising required for producing the alveolopalatal /ɲ/ and the dorsal lowering required for /a/ is expected to interfere with C-to-V coarticulation in /ɲa/ ([Recasens et al., 1997](#)). From a physiological perspective, gestural conflict arises from the contraction of speech muscles which may exert biomechanical forces that cause two antagonistic movements of the same articulator (e.g. tongue body). For example, retroflexion at the front of the tongue (caused by contracting the front superior longitudinal muscle) counteracts the downward bending of the same part of the tongue caused by contracting the front inferior longitudinal muscle.

The hydrostatic nature of the tongue can be a source of gestural conflict which may not only result from motor activation of antagonistic lingual muscles. This concept has implications for the normal production of isolated sounds as well as for coarticulation resistance ([Gick et al., 2013](#)). An example of the effect of tongue hydrostatics on the normal production of isolated sounds comes from high front vowels. The production of vowels such as /i ɪ/ requires contracting the posterior genioglossus muscle, which advances the tongue root and causes the tongue body to be hydrostatically pushed upward or forward. The anterior genioglossus muscle contracts antagonistically at the same time to counteract the hydrostatic force that could push the tongue too far forward. The two antagonistic forces cause raising the tongue body upward, and this lingual configuration is what justifies the [+high] feature of these vowels articulatorily. This example, nonetheless, does not suggest that the tongue body gesture is the product of the root gesture or vice versa. Rather, it suggests that the hydrostatic properties of the tongue set limits on how far different

parts of the tongue can move to produce phonemes. Such hydrostatic constraints are directly relevant in assimilatory pharyngealization. Gestural conflict arising from the hydrostatic nature of the tongue – more specifically, its volume preservation characteristic – has motivated the study in Chapter 5. It examines whether the hydrostatic volume preservation influences the co-occurrence of tongue root advancement in [j] and assimilatory pharyngealization.

2.3.2.2 Coarticulation and strategies for solving gestural conflict. Whether the gestural conflict is between articulatory targets of the same articulator or different articulators, it does not necessarily cause resistance to coarticulation. Rather, it can stimulate other resolution strategies that vary cross-linguistically, dialectally, segmentally and contextually. The same gestural conflict in similar phonetic environments can be resolved differently across different languages and dialects. A given language may stimulate different responses to the same conflict. In general, resolution strategies can be classified in three categories: deletion, transition and compromise/reduction ([Gick et al., 2013, p. 219](#)). Deletion/elision is viewed by some researchers (e.g. [Browman & Goldstein, 1990](#)) as an extreme on a continuum of gestural reduction (thus, leaving unchanged the phonological representation of the utterance). An example of deletion is reported for Yoruba where /r/ is elided when preceding /i/ in order to avoid the gestural conflict at the tongue root, as /r/ is produced with a retraction and /i/ with an advancement ([Gick & Wilson, 2006](#)).

In this section, I will review cross-linguistic evidence for these responses, elaborate on the multiplicity of reconciliation strategies intra-linguistically and review cases of conflict between coupled articulators, as the coupling between tongue root and dorsum will be a crucial element of this dissertation. The latter two concepts, multiplicity of reconciliation and inter-articulator conflict, are particularly relevant to the present study because 1) the gestural conflict that is

expected to interfere with assimilatory pharyngealization at the tongue root arises from either the same articulator or a different articulator, the dorsum; and 2) this conflict leads to multiple reconciliation solutions, namely, resistance to coarticulation and reduction (which is a form of compromise).

2.3.2.2.1 Multiplicity of reconciliation strategies. Cross-linguistic evidence reveals that languages employ different mechanisms for resolving the conflicts between opposing gestural targets. For instance, the conflict between tongue root advancement in a tense vowel/glide and root retraction in a following liquid/uvular elicits various solutions ([Gick & Wilson, 2006](#)). These include an insertion of a transitional schwa between the vowel/glide and the liquid (as in English and Beijing Chinese), laxing the tense vowel (as in Korean) or a combination of both insertion and laxing (as in Nuu-chah-nulth and Chilcotin). In English, for example, when a tense vowel precedes a liquid as in “heel” and “higher”, a schwa-like gesture is inserted and is perceptually salient. Several phonological accounts have been proposed to explain this case of schwa epenthesis ([Lavoie & Cohn, 1999](#); [Orgun, 2001](#)). A prosodic account of this phenomenon was proposed based on the finding that this excrescent schwa is not an inserted vowel or an additional mora because it does not increase syllable duration ([Lavoie & Cohn, 1999](#)). More importantly, based on an ultrasound and acoustic investigation, [Gick and Wilson \(2006\)](#) suggested that the gestural conflict at the tongue root causes a fast-transitional movement of the tongue root within the articulatory (and acoustic) space of a canonical schwa in Chilcotin and Nuu-chah-nulth. This schwa-like transitional gesture takes place as the tongue root moves from its advanced position of the high vowel to the retracted position of the following liquid.

Languages can use different strategies to solve the same gestural conflicts. Nuu-chah-nulth and Chilcotin respond to the conflict between retracted consonants /q s^ʰ/ and the tense vowel /i/

using two different strategies – i.e. transition and compromise – depending on the direction of coarticulation ([Gick & Wilson, 2006](#)). A transitional schwa is inserted in regressive consonant-dependent coarticulation in Nuuchahnulth (e.g. /ci:qci:qa/ realized phonetically as [ci:əqci:əqa] “talking” with an extra [ə] in the transition between [i:] and [q]) and in progressive coarticulation in Chilcotin (e.g. /sʰit/ realized as [sʰəit] “kingfisher”). A compromise strategy is also found in these languages. Since lax vowels require less root advancement than tense vowels, the tense vowel is laxened (lowered) in progressive consonant-dependent coarticulation in Nuuchahnulth (e.g. /qitʃin/ realized as [qetʃin] “louse”) and in regressive coarticulation in Chilcotin (e.g. /niqin/ realized as [neqəin] “we paddled”). What these examples illustrate is that a given gestural conflict can be resolved with more than one method inter- and intra-linguistically and that the direction of coarticulation is sometimes a participatory factor in this reconciliation.

Similarly, a single language can employ more than one method for resolving the same articulatory conflict and this multiplicity of reconciliation strategies cannot always be attributed to contextual factors. In Catalan, when the velar stop /k/ occurs in a sequence of three consecutive consonants /C₁k#C₃/ (where # refers to word boundary), it is prone to either lenition, elision or an increase in its articulatory prominence ([Recasens & Mira, 2015](#)). The same phonetic context triggers either prominence (in about half the data) or reduction/elision (in the other half) contingent on the severity of the lingual constraints. For example, the stop closure of /k/ tends to be elided when preceded by the alveolar fricative /s/, owing to the difficulty of forming a dorsal contact after the generation of a long turbulence for /s/ with a more anterior lingual aperture. Given the acoustic and linguo-palatal contact size which is taken as an index of velar reduction, the stop closure is reduced in the context of sonorants /l r/, which are less lingually constrained. Taken together,

articulatory reduction and elision is conditioned by the lingual constraints of the adjacent obstruents, and to a lesser extent, sonorants.

Also, as documented in [Recasens et al. \(1997\)](#), the gestural antagonism in C-to-V coarticulation may lead either to a more prominent coarticulation or to a greater coarticulation resistance. The increase in C-to-V coarticulatory prominence is particularly common when the source consonant is more lingually constrained than the target vowel. This is based on evidence that, due the high articulatory demands for sustaining frication of [s] which makes this fricative lingually constrained, the adjacent high vowel [i] (as in [si]) is produced with a lower dorsum than in its canonical articulation. In general, these examples suggest that a given articulatory conflict may evoke multiple (and often contradictory) outcomes and that response to gestural antagonism is determined by the articulatory constraints of both the source and target segments ([Recasens et al., 1997](#)).

2.3.2.2.2 Resolving inter-articulator conflict. The coarticulation literature has extensively studied coarticulatory resistance caused by conflicting gestural demands on the same articulator, but there has been little research on gestural conflict between coupled articulators. Independent articulators, such as the velum and larynx, are those that can move freely without interfering with other articulators' targets ([Gick et al., 2013, p. 211](#)). Because of their relative independence, the coarticulation of nasality and voicing, for example, presents less difficulty compared to coarticulation involving more dependent (or coupled) articulators. In contrast, anatomically coupled articulators that are interconnected by muscles or other tissues impose more difficulties on coarticulation, thus giving rise to coarticulation resistance or to other articulatory adaptation mechanisms. The presumed anatomical coupling between the tongue root and dorsum, which might interfere with pharyngealization, will be discussed in the following sections.

Most of the examples so far deal with gestural conflict at a single articulator. There are cases, however, where a conflict arises between different articulators that could be anatomically coupled and functionally dependent. Based on most cases documented in the literature, inter-articulator conflict occurs when a gesture by a given articulator (e.g. advancing the root in palatals) takes place as a consequence of achieving an articulation target with another articulator (e.g. raising the dorsum). This bio-mechanical coupling between two different articulators imposes a constraint on the non-primary articulator (i.e. the tongue root in this example) that could, in one way or another, interfere with coarticulatory root retraction. Inter-articulator conflict stems from either or both the hydrostatic property of the tongue (i.e. volume preservation) and/or the activation of muscles that produce physiologically antagonistic motions (e.g. superior and inferior longitudinal muscles).

Inter-articulator conflict is given a special status in the degree of articulatory constraints (DAC) model and is of a special relevance to the present study as the hypothesis of the resistance to assimilatory pharyngealization is based on a conflict between different, yet functionally dependent, articulators: the tongue root and the dorsum. One such conflict is caused by tongue dorsum raising in Catalan /n/ ([Recasens et al., 1997](#)). Although this tongue raising is just a by-product of primary tongue blade raising, /n/ is specified with a moderate dorsal constraint which, to some extent, constrains the coarticulatory dorsal lowering imposed by adjacent sounds such as /a/. Recasens's linguopalatal contact data reveals that in sequences such as /ana, ina, ani/, the degree of [n] resistance to coarticulatory lowering from [a] is smaller than the resistance of [s ʃ] which are specified for a higher dorsum, and greater than the resistance of [p] which does not involve any antagonistic lingual gesture. Another example where the bio-mechanical dependency between articulators imposes constraints on coarticulation is reported for jaw movement in Catalan

vowels ([Recasens, 2012](#)). Jaw height in vowels is correlated with other articulatory postures such as dorsal raising and oral closing, unlike consonants where jaw height is independent of lingual and labial postures ([Fletcher & Harrington, 1999](#)). Although jaw movement is not the primary articulatory target in vowel production (but is rather associated with the primary lingual/labial articulation), it exhibits some resistance to coarticulation and the degree of its coarticulation resistance is correlated with its height. To illustrate this, the vowel /u/, which involves a high jaw position because of lip closing, exhibits less coarticulatory variability at the jaw than the unrounded [i] which requires a lower jaw position. What these examples show is that a non-primary gesture of the target sound can restrict its coarticulatory variability to some extent. Still unknown, however, is the role of uncontrolled hydrostatic gestures of the source sound (not of the target one) in coarticulation. If, for example, dorsum lowering in pharyngealized consonants is a by-product of tongue root retraction, does this make the antagonistic dorsum raising less effective in restricting assimilation than the antagonistic root advancement? This will be the subject of Chapter 5.

2.4 Coarticulation Models

Several models of coarticulation differing substantially in terms of the basic input unit of speech production have been proposed over the last a few decades. The models proposed the syllable ([Kozhevnikov & Chistovich, 1965](#)), the feature ([Cohn, 1990, 1993](#); [Henke, 1966](#)), and the gesture ([Bell-Berti & Harris, 1981, 1982](#); [Bell-Berti et al., 1995](#); [Keating, 1990](#); [Öhman, 1966, 1967](#); [Recasens et al., 1997](#)) as their primitive input unit. This section discusses both feature- and gesture-based models with a special emphasis on Cohn's feature-based model and Recasens' gesture-based model since they are the most well developed and relevant to the findings of the present study. The exploration of these models will conclude with a brief discussion of their predictions on the assimilatory effect of emphatics.

2.4.1 Feature-based models. The basic premise of these models is that segmental processes such as coarticulation and assimilation are contingent on phonological underspecification, a development of SPE's full specification and binary representation of features in the landmark work of [Chomsky and Halle \(1968\)](#). In contrastive underspecification, as advocated by [Clements \(1985, 1987\)](#) and [Steriade \(1987\)](#), a segment does not inherently have a binary value for a given feature if it does not contrast for this feature and is, thus, classified as being unspecified for the feature.

2.4.1.1 Early feature-based models. One of the early models of this type is the 'dynamic articulatory model' by [Henke \(1966\)](#), known later as the 'look-ahead' model, which maps phonological representations into phonetic targets. More specifically, the model uses as its input a string of phonemes, each associated with attributes that were less abstract than the distinctive phonological features that were dominant at the time. Using this input, every portion of the speech apparatus tries to achieve a phonetic, articulatory target continuously over time. At every instant in time, there is only one dominant phoneme. Importantly, the model "looks ahead" from the current dominant phoneme to the upcoming phonemes and their features and starts anticipating their articulatory configuration provided that their featural demands do not conflict with the demands of the current phoneme. This feature scanning, which is based on binary features, results in categorical assimilation and, hence, it influences the articulation of the current phoneme. However, in cases of featural conflict between different phonemes, the requirements of the current phoneme take precedence and inhibit the influence of future ones. Obviously, this look-ahead mechanism applies only to regressive coarticulation. The model does not assume another mechanism to account for progressive coarticulation. It is programmed at the lower levels – that is, it is considered as a mechanical effect such as physical inertia rather than a cognitive pre-

planning as in regressive coarticulation. The model was further developed by other researchers ([Benguerel & Cowan, 1974](#); [Daniloff & Moll, 1968](#); [Moll & Daniloff, 1971](#)), which took the model further and regarded the ‘look-ahead’ procedure as contingent on feature underspecification.

2.4.1.2 Recent feature-based models. A family of models known as target-and-connection models emerged later ([Keating, 1985, 1988a, 1988b, 1990](#); [Lindblom, Pauli, & Sundberg, 1975](#); [MacNeilage, 1970](#)) and was based on similar principles: phonological underspecification and the mapping from phonological representation into phonetic targets through interpolation.

The target-interpolation model proposed by [Cohn \(1990\)](#), which constitutes the framework of the current study, is a development of Keating’s Window model and is based on observations of a case study of the feature [nasal] in English, Sundanese and French. This phonetic implementation model maps phonology and phonetics by translating discrete phonological features (assigned to segments either underlyingly or via phonological rules) onto phonetic targets specified in time and place. In general, the model postulates that segments that are unspecified underlyingly can undergo either phonological processing or phonetic implementation, with the former taking precedence. That is, these segments are subject to phonological rules such as feature spreading and deletion (if the language has any) within the limits of phonological constraints. In this case, a phonologically assimilated segment must have the same realization of the phonetic properties associated to a feature as a segment underlyingly specified for this feature. Otherwise, the segment is unspecified at the output of phonology and, subsequently, becomes subject to phonetic implementation through interpolation between the surrounding specified segments.

In this model, coarticulation primarily takes place in segments that are unspecified at the output of the phonology and can be accounted for by a phonetic interpolation that linearly connects the phonetic targets of the phonologically specified segments within language-specific phonetic

constraints. With this interpolation, the model assumes that coarticulation is either constant (i.e. plateau-pattern) or changing (i.e. rising or falling cline-pattern) over time. A discussion of the basic principles of the model and an example derivation of nasalization in English using the model follow.

In general, the model uses the phonological representation of segments – as derived by the interaction between feature specifications and phonological rules – as a basis for phonetic implementation. Thus, the model can be summarized into two main components: phonological and phonetic, summarized in (1) and (2), respectively.

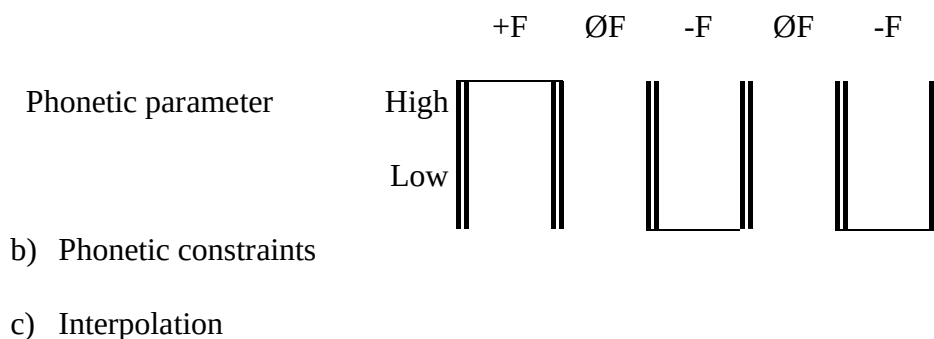
1) Model: Phonology

- a) Underlying representation: features either specified or unspecified
- b) Phonological rules (e.g. feature spreading, deletion, etc.) and constraints
- c) Output of phonology: +F, -F and ØF, where F is the feature undergoing coarticulation

Following [Clements \(1987\)](#) and others, Cohn assumed that segments are unspecified for features that are not contrastive (1a). The phonological (featural) representation is subject to the effect of phonological rules (such as feature spreading) by which segments can acquire new features (1b). Featural specification is also subject to language-specific phonological constraints that trigger underspecification in certain segments: for example, [+continuant] consonants are considered unspecified for the feature Nasal in English (1b). The output of phonology (1c) – i.e. the resultant feature specification after applying the rules and constraints to the underlying representation – is submitted to the phonetic implementation presented in (2).

2) Model: Phonetics

- a) Target assignment:



Phonetic implementation starts with mapping the output of phonology to phonetic targets that are represented over time (rather than being static timeless targets as in other models). As shown in (2a), this mapping applies to specified segments but segments with ØF do not receive targets. These targets are defined on a relevant phonetic parameter, for example, vocal fold vibration for the feature [voice], in which ‘high’ and ‘low’ would refer to the presence and absence of vibration. Phonetic constraints (2b), which only apply to unspecified segments, regulate the phonetic realization of the phonetic parameter in question, for example, the onset of voicing in the case of the feature [voice] and the magnitude of velopharyngeal port closure for the feature [nasal]. Languages may or may not impose such constraints on the phonetic realization of a given feature and these constraints vary cross-linguistically in the sense that a given language may impose a constraint on the onset of the velic closure whereas another language may constrain the degree of closure and not its timing. Finally, within the limits of these constraints, a linear interpolation takes place to connect targets. It can take either a cline or plateau shape as schematized in the first and

second ØF targets, respectively, in (2c). A phonologically unspecified target can receive a plateau-like interpolation for most of its duration only in the absence of phonetic constraints and if the preceding and following specified targets share the same phonetic assignment. It should be emphasized here that segments assimilated by feature spreading leave the phonology specified and, therefore, are not subject to this interpolation. In this case, the segment receives a phonetic target similar to those which are specified underlyingly.

An example derivation of phonetic nasalization in English vowels based on nasal airflow data collected by [Cohn \(1990\)](#) can illustrate the model. Vowels in English are unspecified for the feature [nasal] and there is no phonological rule of anticipatory or carryover vowel nasalization; consequently, they are unspecified at the output of phonology as represented in (3a). Only consonants in the three words are specified and are, therefore, assigned to phonetic targets as in (3b) in which the amplitude of nasal airflow is the phonetic parameter distinguishing targets: [+nasal] consonants are assigned to nasal targets (as indicated by the top horizontal line along the target duration) and [-nasal] consonants are assigned to oral targets (as indicated by the low horizontal line along its duration). Phonetic constraints are then imposed: voiced stops [b d] must be oral at the stop release (as represented with the little square at the bottom-right corner of [b] and [d] phonetic targets in 3c) and the magnitude of vowel nasalization must be less than that of [+nasal] consonants (as vowels phonetic target is represented in 3c with a lower nasalization value, i.e. lower horizontal line, than in the triggering nasals). The latter is based on laboratory evidence in [Kent, Carney, and Severeid \(1974\)](#) which reported a limited velum lowering in vowels compared to consonants. The model assumes a rising cline of anticipatory nasalization in [bin] and a falling cline of carryover nasalization in [nid] within the limits of the phonetic constraints on vowels as in (3d). Simultaneous anticipatory and carryover nasalization in [mɛn] takes a plateau-

like interpolation, constant magnitude of nasalization throughout the target duration but with a lower amplitude than in the targets of [m] and [n] given the phonetic constraints on vowels. (It is worth noting here that Cohn used English vowel nasalization as an example of phonetic implementation whereas studies like [Solé \(1992\)](#) argued that it is phonological. Solé's claim was based on her findings that the nasalized portion of a vowel varied proportionally to the speaking rate, reflecting a preserved perceptual difference between oral and nasalized vowels in different speech rates).

3) Example derivation of English vowel nasalization using Cohn's model

a) Output of phonology:

b	i	n	n	i	d	m	ε	n
-N		+N	+N		-N	+N		+N

b) Assignment of phonetic targets:

Nasal airflow	High	b	i	n	n	i	d	m	ε	n
	Low									

c) Phonetic constraints:

b	i	n	n	i	d	m	ε	n
.					.			

d) Interpolation:

b	i	n	n	i	d	m	e	n
.					.			

2.4.2 Gesture-based models. Several models such as the ‘time-locked’ model ([Bell-Berti & Harris, 1981, 1982](#); [Bell-Berti et al., 1995](#)) consider the articulatory gesture to be the primary input of speech production. They were devised to account for the dynamic nature and temporal extent of coarticulation, which was not sufficiently addressed in the earlier model by Henke. For example, the central principle of the time-locked model is that the duration and start time of an articulatory movement differ across different articulators and that the duration of anticipation is independent of the length of preceding phones. As the temporal aspect of coarticulation is not central to the present study, these models (although gesture-based) will not be discussed further. I will instead focus on gesture-based models that address coarticulation resistance, which is more crucial to this study.

The concept of coarticulation resistance, one of several approaches proposed in the literature to account for the (inter)relationships between invariant speech sounds, first emerged in 1960s ([Öhman, 1966, 1967](#)) and was later developed in the 1970s ([Bladon & Al-Bamerni, 1976](#); [Bladon & Nolan, 1977](#)). It inspired the more recent degree of articulatory constraints (DAC) model ([Farnetani & Recasens, 2010](#); [Recasens et al., 1997](#)) which accounts for diachronic processes originating from coarticulation in Romance languages, such as assimilation, segmental insertion and elision ([Recasens, 2014](#)).

2.4.2.1 Early models of coarticulation resistance. In his cross-linguistic comparison of coarticulation, [Öhman \(1966\)](#) argued that articulatory demands interfere with coarticulation. Based on an examination of regressive vowel- and consonant-dependent coarticulation in VCV sequences from English, Swedish and Russian, Öhman’s model regarded the tongue as consisting of three articulators that operate independently despite sharing some muscles. English and Swedish stop consonants are produced exclusively by the apical and dorsal articulators and not the tongue body,

which leaves them more freedom to coarticulate with adjacent vowels. This does not apply to Russian stops where the production of palatalization requires employing the tongue body in addition to the apical and dorsal articulators. Therefore, Öhman proposed that although the three articulators are controlled independently, phonetic restrictions in some languages such as Russian may override this independent control and impose restrictions on coarticulation. This theoretical conceptualization was then implemented in a numerical model which attempts to indirectly predict coarticulation resistance based on possible vocal tract shapes formed by the interaction of the three articulators ([Öhman, 1967](#)).

Building on Öhman's model, Bladon and his colleague proposed that coarticulation resistance, which inhibits coarticulation, is contingent on the articulatory requirements of the segment undergoing coarticulation and that coarticulation resistance is allophone-based because allophonic variants resist coarticulation to different degrees ([Bladon & Al-Bamerni, 1976](#); [Bladon & Nolan, 1977](#)). For instance, the variants of /l/ in British Received Pronunciation exhibited varying degrees of resistance, where clear [l] is consistently the most contextually variable allophone, syllabic [ɫ] the most resistant, and non-syllabic [ɫ̥] falling somewhere in between. The hierarchical organization of sounds in terms of their coarticulation resistance is based on rules that are language-particular and, in some cases, quasi-universal. Also, allophones' resistance is independent of the property undergoing coarticulation and of the direction of coarticulation. Moreover, assignment of coarticulation resistance values is idiolectally-based, given that the extent of vowel-dependent coarticulation on nasals has a speaker-identifying function ([Su, Li, & Fu, 1974](#)). The model also assumes that a resistant segment such as /s/ can also impede the coarticulatory effect from spreading to neighboring segments, based on laboratory evidence in [Amerman, Daniloff, and Moll \(1970\)](#).

2.4.2.2 The degree of articulatory constraints (DAC) model. Inspired by Bladon and Öhman's models, the DAC model was proposed to account for lingual coarticulation and quantify the differences across segments in terms of their coarticulatory resistance and dominance ([Farnetani & Recasens, 2010](#); [Recasens, 1989](#); [Recasens & Espinosa, 2009](#); [Recasens & Mira, 2015](#); [Recasens & Pallarès, 1999](#); [Recasens et al., 1997](#)). The model has been recently extended to encompass other sound changes and processes that may originate from coarticulation such as segmental insertion, epenthesis, assimilation and dissimilation ([Recasens, 2014](#)); however, the discussion here focuses on the operating principles of the DAC as they relate to coarticulation resistance specifically.

The basic principle of the DAC model is that the articulatory requirements of the segment subject to coarticulation (including, but not limited to, the gestures, lingual shape and place and manner of articulation) can contribute to its coarticulatory variability in terms of magnitude, temporal extent and direction. Place of articulation plays a special role as the degree of articulatory constraints imposed on a segment increases with the active participation of the lips and/or a lingual region in the formation of a constriction. For example, labials are constrained at the lips and, therefore, more resistant to labial coarticulation than lingual consonants such as /l/; nonetheless, they are not constrained at the tongue root and should be sensitive to tongue root retraction. This means that labials should be very sensitive to pharyngealization since their production does not involve any antagonistic gesture. An important implication of this principle is that primary articulatory gestures induce greater constraints on coarticulation than the uncontrolled tongue gestures which are caused by bio-mechanical forces. For example, since the dorsum involvement in the formation of the constriction in the Catalan alveopalatals [ʃ tʃ] is the primary gesture, these consonants are more constrained at the dorsum than [t n]. The dorsum movement in the latter, [t

n], is only a consequence of the movement of the front primary articulator due to anatomical coupling between the two articulators ([Recasens, 2014](#)). Similar evidence from [Recasens and Espinosa \(2009\)](#) supports this claim: coarticulatory variability of /n l s/ at the dorsum, which is not an active articulator, is greater than at their primary articulator, tongue tip and blade. If this claim holds true for the tongue root, it is expected that segments with velar raising as a primary gesture (e.g. [k]) shall be more resistant to assimilatory pharyngealization than segments with secondary velar raising (e.g. [n]). This study, in fact, does not directly compare these two consonants to each other and only assumes that since the articulatory constraints on the posterior area for [t n] are negligible, they can be used to test other hypotheses such as the effect of phonemic contrast.

Manner can also impose constraints on coarticulation, as illustrated by evidence of greater resistance of frication than laterality ([Recasens & Espinosa, 2009](#)) and greater resistance by the trill [r] than the tap [ɾ] ([Recasens & Pallarès, 1999](#)). Nevertheless, the model states that constraints by manner requirements are less strict than place restrictions, and the small differences in resistance between palatals of different manners accords with this claim ([Recasens & Rodríguez, 2016](#)).

The factors explained so far (i.e. place and manner restrictions and primary versus secondary articulation) have led to the formulation of the DAC scale, a quantification of the difference across consonants and vowels in the degree of their lingual articulatory constraints. The scale was formulated to account for dorsum constraints (because it was based on EPG data which allowed imaging the palatal zone only) although the model itself incorporates other lingual regions (e.g. the palatal, velar and pharyngeal zones) based on more recent ultrasound evidence of these regions ([Recasens & Rodríguez, 2016, p. 72](#)). The DAC values predict the degree of articulatory constraints at the tongue dorsum of the segment they are bound to ([Recasens, 2014, pp. 1-8](#)). On

the top of the scale are the most dorsally constrained segments which require high articulatory precision and are given a DAC value of 3: (alveolo)palatal consonants and vowels, velars, the trill [r] and dark [ɫ]. At the bottom of the scale are the bilabials and the schwa which are the least lingually constrained, the least resistant to coarticulation and, thus, given a value of 1. Between these two extremes lies a set of consonants and vowels where the dorsal activity is not primary but rather secondary: dentals, the low vowel [a] and alveolars except [ɫ r].

According to the DAC model, there are other factors that may not contribute to coarticulation resistance (and, therefore, are not accounted for in the DAC hierarchy) such as syllabic position. Since consonants in syllable onset were not found to be articulatorily more prominent than coda consonants (with the exception of an increased consonant duration), syllable position by itself does not impact consonant resistance ([Recasens & Rodríguez, 2018](#)). Rather, resistance depends on the articulatory constraints of the consonant occupying the onset versus coda position as well as the dominance of anticipatory versus carryover effects associated with the consonant. For example, less constrained consonants /t n l/ are equally variable in both syllable positions whereas other consonant such as /p s r ɲ/ are more variable at the syllable onset than at the coda and /k/ has the opposite pattern ([Iskarous & Kavitskaya, 2010](#)). However, since the effect of syllable position on coarticulation is not consistent across all consonants, languages and coarticulatory phenomena, this factor was controlled for in the present study to avoid any potential confounding effects.

Several other general principles underlie the DAC model and relate to the present study. Namely, the model predicts a positive correlation between the coarticulatory resistance and dominance (“aggressiveness”) of a given segment and that both are determined by its articulatory constraints. This prediction is true for both consonants and vowels; however, it holds for the tongue

body rather than the tongue front ([Recasens & Rodríguez, 2016, p. 73](#)) because the tongue body activity spans a greater duration than the activity at the tongue front, and this increases the likelihood of extending the tongue body movement to neighboring segments. Another principle is that consonants and vowels favor one coarticulation direction over the other and that the prevalence of either direction depends on the temporal characteristics of consonant production. To illustrate this, the anticipatory effect of dark [ɪ] in Catalan is more prominent than its carryover effect because the tongue dorsum lowering and retraction for [ɪ] (which are initiated before tongue tip raising) need to take place during the preceding vowel ([Recasens, 2014, p. 9](#); [Recasens et al., 1997, p. 559](#)).

Furthermore, the DAC model predicts a correlation between the magnitude and temporal extent of coarticulation, both of which are contingent on the articulatory constraints of a segment. According to [Recasens et al. \(1997\)](#), larger coarticulatory effects are temporally longer, e.g. C-to-V effect in /li/, and smaller effects are temporally shorter, e.g. C-to-V effect in /la/. The model predicts that since /l/ is highly constrained, its coarticulatory effect on the adjacent vowel is more prominent in the case of gestural antagonism, as in /li/, than in the case of gestural compatibility, as in /la/. Two articulatory gestures contribute to the antagonism in /li/: one at the dorsum and another at the root. The dorsum conflict, according to Recasens, arises because dark /l/ is produced with dorsum lowering and backing whereas /i/ is produced with dorsum raising and fronting. This antagonism is absent in /la/ where the two segments are produced with dorsum lowering. The other antagonism, as documented in English ([Gick & Wilson, 2006](#)), is at the tongue root as /l/ requires root retraction and /i/ root advancement.

The last principle to be mentioned here is that resistance and dominance of C-to-V coarticulation are conditioned by the relative DAC specification of the two segments undergoing

coarticulation (i.e. source and target segments) as well as the antagonism or compatibility of their lingual shapes. When the two segments are specified for the same DAC value, the model predicts a large effect in the case of articulatory antagonism (e.g. /li/) and a small effect in the case of articulatory compatibility (e.g. /ni/). Coarticulation prominence increases when the DAC specification is higher for the triggering consonant than the context vowel (especially with gestural antagonism) and decreases when the vowel's DAC is higher than the source consonant's DAC.

2.4.2.3 Comparison of coarticulation resistance models and preference for DAC model.

The gestural models of coarticulation discussed so far adhere to the basic tenet that the extent to which speech sounds retain the articulatory properties of their underlying form is conditioned by their resistance to coarticulatory variability and that this resistance is, in turn, contingent on their articulatory requirements. These models also share the concept of a quantified hierarchical resistance across consonants and vowels. For example, the three models classify the alveolar fricatives /s z/ as being highly resistant to vowel-dependent coarticulation compared to other consonants. Another commonality across these approaches is that none of them claims that articulatory constraints are entirely universal: Bladon's coarticulation resistance is language-particular or quasi-universal, while Recasens's DAC specifications are dialect-specific. For example, articulatory difference in the production of [s], which is articulated with a more forward constriction in the Valencian dialect of Catalan than in the Eastern dialect, causes the two dialects to differ in terms of the resistance of adjacent vowels to the coarticulatory effect of [s] ([Recasens, 2014, p. 4](#)). Moreover, although the DAC model is phoneme-based, its specifications can be allophone-based, like Bladon's model, particularly in cases when different allophones of the same phoneme make different predictions (e.g. effect of [r] vs. [ɾ] and [l] vs. [ɫ]).

Nevertheless, due to many substantial differences between the three models, the DAC model in particular will be the main focus in this dissertation. Firstly, the DAC model accounts for several aspects of coarticulation that are relevant to this research but are not addressed by the earlier models. For example, unlike Bladon's model, which accounts for coarticulation at the anterior lingual region, the DAC model encompasses the posterior tongue zones, which are particularly important to this study, in addition to front articulators (e.g. the jaws). A second reason is that the basic primitive that governed coarticulation resistance in previous models was the articulatory gesture whereas the DAC model relies on a wider array of articulatory requirements, such as the manner of articulation and the degree of lingual constriction, in addition to articulatory gestures. The DAC model also addresses several other factors that influence coarticulation in one way or another such as syllable position, and factors which the other models predicted to be non-influential (e.g. the direction of coarticulation as some consonants, according to [Recasens et al. \(1997\)](#), favor regressive coarticulation over the progressive one, and vice versa). Although these particular factors are not central to the present study, they were taken into consideration in designing it. These reasons make the DAC model more relevant to this research.

However, it should be noted here that this study will not try to develop a DAC scale to quantify differences between segments and their resistance to emphasis. Formulation of such a scale requires articulation and acoustic data that must be elicited in a wide range of coarticulation conditions that span several segments, phonetic contexts and many other forms of variation. This study will rely on the explanatory mechanisms put forward in DAC models, without trying to develop specific quantification of gestural resistance.

2.5 Predictions of Emphasis Spread

A core issue in this dissertation is the impact of the aforementioned phonological and physiological constraints on emphasis spread and how they relate to the postulates of the two coarticulation models explained earlier. Since these constraints and the models are conceptually related, two groups of predictions emerge: the predicted effects of phonological constraints are derived from Cohn's feature-based model; likewise, the predicted gestural antagonism to emphasis is derived from Recasens' gesture-based model. Therefore, I present here the predictions derived from both models together while emphasizing the cases where the two models lead to contradictory predictions about sounds' sensitivity to emphasis spread. That is, a given sound can be phonologically constrained and, thus, predicted not to undergo assimilation (according to the phonological constraints and Cohn's model) but physiologically unconstrained and, thus, predicted to be able to assimilate freely (according to the physiological constraints and Recasens' model). Table 2.1 summarizes the predicted effect of pharyngealization on all consonants based on the presence of phonological constraints (i.e. phoneme contrast) and physiological ones (i.e. gestural antagonism). Detailed predictions are discussed in the remainder of this section.

Table 2.1

Summary of predicted effects of pharyngealization on all consonants according to the type of constraints on the consonant

Place of articulation	Consonant	Predictions based on constraint type	
		Phonological constraints	Physiological constraints
Labial	b	Unconstrained	Unconstrained
	m	Unconstrained	Unconstrained
	f	Unconstrained	Unconstrained
Dental	t	Constrained	Unconstrained
	d	Constrained*	Unconstrained

Place of articulation	Consonant	Predictions based on constraint type	
		Phonological constraints	Physiological constraints
	n	Unconstrained	Unconstrained
	θ	Unconstrained	Unconstrained
	ð	Constrained	Unconstrained
Alveolar	r	Unconstrained	Unconstrained
	s	Constrained	Unconstrained
	z	Unconstrained	Unconstrained
	ʃ	Unconstrained	Highly constrained
	ʤ	Unconstrained	Highly constrained
	l	Unconstrained	Unconstrained
Palatal	j	Unconstrained	Highly constrained
Velar	k	Unconstrained	Moderately constrained
	g	Unconstrained	Moderately constrained
	x	Unconstrained	Moderately constrained
	ɣ	Unconstrained	Moderately constrained
	w	Unconstrained	Moderately constrained
Uvular	q	Unconstrained	Moderately constrained
Pharyngeal	ħ	Unconstrained	Unconstrained
	ʕ	Unconstrained	Unconstrained
Glottal	ʔ	Unconstrained	Unconstrained
	h	Unconstrained	Unconstrained

* If phonemic contrast is what restricts coarticulation, [d] should be resistant. If allophonic contrast is what restricts it, [d] should not be resistant.

2.5.1 Predicted effects of phonological constraints and the feature-based model.

Assimilatory effects can be categorical or gradient depending on whether the factors at play are phonological or phonetic, respectively. Hence, phonemic contrast makes binary predictions in the

sense that a sound is either constrained or not, which is consistent with the phonological-level processes assumed by Cohn.

The emphatic phonemes in all Arabic varieties are /tˤ sˤ dˤ ðˤ/ and contrast with /t s d ð/, respectively; however, the allophonic realization of the voiced /dˤ ðˤ/ differs across dialects. In the Eastern Peninsular Arabic, both of these consonants /dˤ ðˤ/ are realized as [ðˤ] except in a few sociolects³. This means that this dialect contrasts emphatic [ðˤ] and its plain counterparts [d ð] (instead of a distinction between [dˤ ðˤ] and [d ð]). Emphatic consonants in Arabic are represented formally with multiple distinctive features including:

- [+RTR] under the Pharyngeal node, signaling the retraction of the tongue root or a constriction in the pharynx and, thus, referring to two articulators, the tongue root and larynx ([Halle, 1995](#); [Rose, 1996](#));
- [+pharyngeal] under the Pharyngeal node, characterizing the ‘region’ of the vocal tract spanning from the larynx to the oropharynx inclusively (rather than a single articulator) and thus encompassing three articulators, the larynx, tongue root and body ([McCarthy, 1991, 1994](#));
- [+back] under the Tongue Body node, signaling the backing of the tongue dorsum in forming the secondary articulation, which is compatible with the view that emphatics are uvularized rather than pharyngealized ([Bin-Muqbil, 2006](#));
- [+flat] based on the acoustically based feature system, the Preliminaries of [Jakobson, Fant, and Halle \(1951\)](#).

³ Most studies in Arabic phonology, including the ones cited in this section, agree that /ðˤ/ and /dˤ/ are two distinct phonemes that surface as [ðˤ] in many dialects. However, it is possible that this distinction is sustained due to the influence of Modern Standard Arabic in which many of the speakers are fluent in and which contrasts /ðˤ/~dˤ/.

2.5.1.1 Predictions based on phonological contrast. With the exception of [t s d ð], emphasis should spread to all phonemes and allophones, including the coronals [n l] addressed in this study, because they are not contrastive for the feature undergoing assimilation. On the other hand, due to being contrastive in emphasis, the coronals [t s d ð] are predicted to react differently to assimilation – and more precisely, emphasis spread should be restrained to preserve phonemic distinctiveness. The predicted assimilatory pattern of /t s/ should be the same regardless of the surface or underlying distributions. However, the phonetic realization of [ð^s] could be conditioned by its contrastive relation with /d ð/, or only on its surface opposition to [ð] ([Mok, 2012](#); [Mok & Hawkins, 2004](#)).

If phonemic contrast between /d ð/ and /d^s ð^s/ is what matters, both plain consonants should resist assimilation in order to avoid confusion with their phonemic counterparts. On the other hand, if allophonic contrast is what affects assimilation, only [ð] should resist assimilation. The other allophone [d] would be predicted to assimilate more freely because assimilation would produce the sound [d^s] which is not an allophone that is otherwise present in this dialect. In other words, assimilation of [d] does not undermine perceptual and allophonic categorization. Unfortunately, given the low frequency of /ð^s/ and the discrepancy between its surface and underlying form in the dialect under study, it turned out to be nearly impossible to come up with a well-controlled wordlist testing the behavior of these consonants. For this reason, investigating the potential effect of phoneme contrast in this dissertation will remain limited to voiceless [t s].

2.5.1.2 Predictions of the feature-based model. Given that emphasis is characterized with multiple features, the predictions of any feature-based model are largely determined by what feature is being spread. This raises the issue of the discrepancy between feature-based predictions in the sense that some emphatics features can spread while others cannot, depending on whether

the affected segment is unspecified for that feature or not. Given this discrepancy, the present study adopts the feature [RTR] for characterizing emphatic consonants and the predictions here are discussed in light of this feature. The rationale behind selecting this feature is twofold. First, it is the most agreed upon feature for emphatics in the literature ([Halle, 1995](#); [Halle et al., 2000](#); [McOmber, 1996](#); [Shahin, 1996, 1997, 2011](#)). Second, as will be argued in §3.5, it is the feature most compatible with the laboratory data obtained in this study in terms of the involvement of the tongue root in constriction formation; there is much less evidence that the tongue body is the primary articulator responsible for emphasis (which would be entailed by the feature [back]). Furthermore, the predictions discussed here assume contrastive underspecification ([Clements, 1987](#); [Steriade, 1987](#)), which postulates that a segment is unspecified for a certain feature unless it contrasts with another segment in terms of that feature, an assumption underlying the target-interpolation model by [Cohn \(1990\)](#) that is used to model the findings of this study.

Taking these assumptions into consideration and following Cohn's model, Arabic consonants should react differently to the assimilatory effect of emphatics depending on their (under)specification for the feature [RTR] as proposed by [McCarthy \(1991\)](#). According to the model, segments unspecified for [RTR] should be subject to the effect of the feature trigger at the phonological or phonetic level, and this applies to all Arabic consonants that are underlyingly unspecified for this feature (all but /tˤ sˤ dˤ ðˤ/, their plain cognates /t s d ð/ and other gutturals /q χ ʁ ħ ʕ/).

To illustrate this, the model predicts that since the plain coronals /t s d ð/ are underlyingly specified as [-RTR], as they contrast with the emphatic ones /tˤ sˤ dˤ ðˤ/ bearing [+RTR], they cannot adopt [+RTR] phonetically via interpolation. Therefore, even if these plain coronals co-exist with a [+RTR] trigger, they are mapped to a plain phonetic target with no tongue root

retraction (i.e. articulatory root retraction being the phonetic parameter for [RTR]). On the same grounds, being underlyingly specified as [+RTR], emphatic consonants /tˤ sˤ dˤ ðˤ/, uvulars /q χ ʁ/ and pharyngeals /ħ ʕ/ are assigned phonetic targets with a retracted tongue root. This assignment takes place at the phonological level and not via phonetic implementation. Other consonants, since they are underlyingly unspecified for this feature, may either acquire the feature phonologically or some degree of retraction phonetically via interpolation. In the former case, acquiring [+RTR] from an adjacent triggering consonant by feature spreading would mean that the phonologically assimilated consonant is assigned to a phonetic target with the same magnitude of tongue root retraction as in the [+RTR] trigger and that its realization must be constant throughout the entire duration of the target. Otherwise, if no feature spreading takes place, the unspecified consonant would be subject to phonetic implementation. Hence, it would not be mapped to any phonetic target at the start of phonetic implementation but would take on a gradient, cline-like tongue root retraction from the phonetic target of the [+RTR] trigger. The magnitude of the tongue root retraction is predicted to either linearly decrease or increase depending on the direction of assimilation: if it is progressive, the retraction should linearly decrease throughout the target duration; if it is regressive, it should linearly increase. The only case in which the phonologically unspecified consonant would be realized with plateau-like target is when it is surrounded by two phonologically [+RTR] consonants (i.e. similar to the derivation of /men/ in (3) above but with greater value of root retraction). Given that emphasis spread on consonants has never been studied before, a plateau-like interpolation reflecting the absence of phonetic constraints on the extent of tongue root retraction on assimilated consonants would be expected.

2.5.2 Predicted effects of physiological constraints and gesture-based model. Since this research investigates the role of gestural conflict in emphasis spread, it is important to understand

what stimulates this antagonism from a physiological and phonetic perspective. Knowledge of the possible muscular structures that are primarily responsible for the pharyngeal articulation helps predict potential sources of gestural conflict which, in turn, can influence emphasis spread in some way.

There are two possible scenarios for producing emphatics: they can be uvularized by raising and retracting the dorsum ([Al-Solami, 2017](#); [Zawaydeh, 1997](#)) or pharyngealized by retracting the tongue root ([F. Al-Tamimi et al., 2009](#); [Delattre, 1972](#); [Laufer & Baer, 1988](#)). Each of the two proposals has its own consequences in terms of which gestures should be antagonistic to emphasis and what phonemes should be incompatible with emphasis spread. Since the uvularization and pharyngealization account of emphasis imply different tongue shapes, some sounds that are gesturally antagonistic to uvularization can be compatible with pharyngealization, and vice versa.

Given all of this and since the laboratory findings of this dissertation (§3.4.1) yielded evidence supporting the pharyngealization proposal (rather than uvularization one), I outline here the predictions potential gestural antagonism to pharyngealization exclusively.

2.5.2.1 Predictions based on physiological constraints. Predicting antagonism is more complex when considering emphatics as pharyngealized, rather than uvularized. What is problematic here is that the horizontal pharyngeal constriction may originate from more than one possible muscular contraction – the hyoglossus muscle or the middle pharyngeal constrictor – and that such muscular inputs have different consequences on assimilation and gestural antagonism ([Gick et al., 2013, pp. 160-180](#)).

First, contracting the middle pharyngeal constrictor pulls the posterior pharyngeal wall forward. If this muscle is responsible for the pharyngeal constriction, it is basically the posterior

and not the anterior wall (tongue root) that would move to constrict the pharynx. In this case, I would expect little or no interference with other tongue movement because the pharynx is independent of the lingual activity; therefore, pharyngealization should be compatible with all consonants. It should be noted here that the ultrasound imaging used in this study does not reveal the movement of the posterior pharyngeal wall and, therefore can neither verify nor eliminate the role of this part of the vocal tract – either as a sole articulator or in coordination with the tongue root. However, what the present study can tell is whether or not the anterior pharyngeal wall is involved in the production of pharyngealization.

The horizontal pharyngeal constriction could alternatively be produced by contracting the hyoglossus muscle, which would give rise to a potential gestural antagonism. Contraction of this muscle results in a backward displacement of the tongue root and, by virtue of the bio-mechanics and volume-preservation of the tongue, lowers down the tongue body. The resulting tongue shape should be quite different from the tongue shapes reported in the studies supporting uvularization (e.g. [Al-Solami, 2017](#)) where the tongue body is raised and should therefore predict a pattern of gestural conflict opposite to that based on uvularization. That is, if pharyngealization is primarily realized by means of root retraction, gestural conflict should arise from dorsal raising, tongue root advancement, or both, yielding a set of predictions completely opposite to those based on the uvularization proposal where the conflict arises from dorsal lowering (and not raising). The most constrained sounds would be those produced with both dorsal raising and root advancement, such as [ʃ ʒ j i]. Less constrained sounds would be those produced with dorsal raising but without root advancement, like the high vowel /u/, the velars [k g x] and the uvulars [q χ]. The sounds that do not involve active dorsal raising (such as [b m f θ ð a æ n l]) should be unconstrained at the dorsum and root and, thus, articulatorily compatible with pharyngealization.

2.5.2.2 Predictions of a gesture-based model. Based on phonetic factors such as articulatory constraints and gestural antagonism, the DAC model proposes a gradient sensitivity/resistance to coarticulation. Some of the factors that contribute to this gradient effect – such as manner of articulation and the temporal extent and direction of coarticulation – are controlled for in the present study and, therefore, not discussed here. The other factors – namely, place of articulation and active versus secondary articulation – are expected to affect coarticulation in this study and are elaborated on here. It must be noted, though, that given the scarcity of articulation studies on Arabic consonants and vowels, the discussion is speculative to some extent.

A consonant's sensitivity to assimilatory pharyngealization is defined according to the involvement of two articulators in its production: the tongue root and dorsum, which are responsible for producing pharyngealization. Conflicting gestures at the tongue root should constrain assimilation to a greater extent than at the dorsum because the root retraction is the active gesture of pharyngealization whereas dorsum lowering is caused by the volume preservation characteristic of the tongue.

Non-lingual consonants such as the bilabials /b m/ and labiodental /f/ are the least constrained at the posterior region of the tongue and, therefore, should be free to assimilate with emphatic consonants. Three other categories of consonants involving a lingual activity should be less sensitive to assimilation: one of them with a tongue root advancement and two with a dorsal raising (secondary or primary). The first category, consonants with an advanced tongue root such as [j ʃ ʒ dʒ], is the most articulatorily constrained and, consequently, should be highly unaffected by pharyngealization. The second category comprises consonants such as velars and uvulars in which the high dorsum is the active articulator. Because of this dorsal constraint, they should be less sensitive to assimilation. The third one is consonants with a secondary dorsal raising stemming

from a bio-mechanical dependency with the tongue crown (e.g. dental /θ ð/ and alveolar /t d n s z l/, in which the dorsum height is mainly due to the tongue lamina raising). This category should be more prone to assimilatory pharyngealization than consonants with primary dorsal raising (e.g. the velar [k g x] and uvular [χ]). Among these categories of phones, the degree of articulatory constraint is predicted to be positively correlated with the degree of gestural antagonism. That is, the higher the dorsum and the more constrained the consonant, the less sensitive to pharyngealization it should be.

These consonants may react differently to assimilation. They may adopt either of two reconciliation strategies reported in the literature to account for the conflict with assimilation: resistance or compromise. In the first case, the articulatory demands of the affected consonant would take precedence over those of assimilation in order to sustain the canonical articulatory configuration of the consonant. Otherwise, the conflict may result in a compromise to accommodate the assimilation demands, in which case it can lead to less dorsal raising or less tongue root advancement. Nonetheless, the reduction of the dorsal or root gesture may allow the consonant to exhibit assimilatory root retraction to a limited extent because of the volume-preservation of the tongue. In this case, the degree of articulatory reduction would be positively correlated with the magnitude of assimilation at the tongue root.

In fact, the DAC model claims that linguistically constrained sounds respond to articulatory conflict by resisting assimilation and does not postulate articulatory reduction as an alternative solution. Nonetheless, it was reported by ([Recasens & Mira, 2015](#)) that the closure duration of /k/ is reduced due to the manner of articulation of the following obstruent. Given this, there is no reason to exclude reduction (compromise) as a consequence of the conflicting demands of pharyngealization.

Importantly, there are two sources of discrepancy between the predictions of Cohn's model and the DAC model in relation to emphasis spread. First, the DAC model predicts a gradient sensitivity to assimilation which is correlated with the degree of articulatory constraints. Cohn's model, on the other hand, proposes either a categorical or a gradient effect depending on whether the effect is derived through feature spreading at the phonological level or through phonetic implementation. Furthermore, a given consonant could be predicted to undergo assimilation by one of the models, but not the other. This is particularly evident in the case of the feature [+RTR]. According to Cohn's model, all consonants which are unspecified for [RTR] (i.e. all except gutturals, emphatic consonants and their plain cognates) should assimilate freely. In contrast, the DAC model predicts that some of these consonants (such as velars and advanced-root consonants) should not exhibit any assimilation.

CHAPTER 3

Is Emphasis Spread Categorical?

3.1 Introduction

An adequate characterization of emphasis requires exploring the articulatory correlates of the secondary articulation of underlyingly emphatic consonants and the exact nature of their contextual effects on neighboring sounds, two areas of continuing controversy among linguists. As explained in Chapter 2, based on several instrumental studies that have examined the nature of the secondary constriction of emphatics, there seems to be variation in the articulation of the secondary constriction: the key articulator can be the dorsum, which has been claimed to be raised towards the uvula and tilted backward, or the tongue root, which has been claimed to retract backwards towards the posterior pharyngeal wall. Other articulatory gestures (tongue body depression, epiglottal retraction, upper/lower pharyngeal constriction, laryngeal elevation) are physiologically contingent on whether the dorsum is raised or the tongue root retracted ([Esling, 1996, 2005](#)). What articulator (i.e. the dorsum versus the root) is actively responsible for articulating these consonants has motivated a debate on whether they are uvularized ([Al-Ani, 1970](#); [Al-Solami, 2017](#); [Bin-Muqbil, 2006](#); [Ghazeli, 1977](#); [Zawaydeh, 1997, 1998, 1999](#); [Zawaydeh & de Jong, 2011](#)) or pharyngealized ([F. Al-Tamimi & Heselwood, 2011](#); [J. Al-Tamimi, 2017](#); [Ali & Daniloff, 1972](#); [Altairi et al., 2017](#); [Hassan & Esling, 2007, 2011](#); [Laufer & Baer, 1988](#); [Laufer & Condax, 1981](#); [Shar & Ingram, 2011](#); [Zeroual & Clements, 2015](#); [Zeroual et al., 2011](#)). This

uvularization-pharyngealization distinction has been attributed to several factors including, but not exclusively, dialect, gender and other sociolinguistic factors ([Khatab et al., 2006](#)).

The other area lacking unanimity is whether the contextual effect of these consonants is assimilatory feature spreading ([S. Davis, 1993, 1995](#); [Herzallah, 1990](#); [Shahin, 1996, 1997](#); [Younes, 1994](#)) or coarticulation ([Embarki, Ouni, Yeou, et al., 2011](#); [Jongman et al., 2011](#)). None of these studies have explicitly justified their terminological choices, suggesting that this conceptualization is arbitrary rather than being empirically grounded. These two aspects of pharyngealization are the subject of the ultrasound experiment reported in this chapter.

3.2 Research Questions and Hypotheses

The purpose of this chapter is to probe the two aforementioned aspects of emphasis in Eastern Peninsular Arabic. First, the position of the dorsum and the tongue root in phonemically pharyngealized consonants /tʰ sʰ/ in comparison to their position while producing their plain counterpart /t s/ is crucial evidence for the characterization of /sʰ tʰ/ as uvularized or pharyngealized. Second, the study is also concerned with the categorical vs. gradient effect of pharyngealized consonants on the preceding sounds as seen in ultrasound imaging and as an indicator of either assimilation or coarticulation. The regressive contextual effect is examined on consonants exclusively (C-to-C effect), an area where the present study diverges from the existing literature on emphasis spread, which is limited to C-to-V effects. In order to achieve these goals, this study addresses the following two questions.

3.2.1 Uvularization vs. pharyngealization. A question that is fundamental to all subsequent experiments in this thesis is whether these speakers' emphatic consonants are articulated with uvularization or pharyngealization. If it is uvularization, the dorsum must be actively raised and retracted towards the uvula (although the uvula is not visible in the ultrasound

images). As a consequence of this diagonal movement of the dorsum, the tongue body must also be raised and tilted backward. The role of the tongue root is predicted to be insignificant and if it exists, it should be limited to a slight retraction due to its anatomical coupling with the dorsum.

However, a completely different tongue shape should be observed if emphatics are produced with pharyngealization. The key articulator would be the tongue root, which is predicted to be extremely retracted against the posterior pharyngeal wall, a movement that would provoke other lingual movements. Two of these concomitant movements would be the depression and retraction of the dorsum, and the lowering of the tongue body. In brief, both the uvularization and pharyngealization hypotheses predict a dorsal retraction, but dorsal raising/lowering and tongue root retraction should provide us with the evidence needed to decide whether it is uvularization or pharyngealization.

3.2.2 Categorical vs. gradient emphasis. A central question to the present study is whether the contextual effect of emphatics on the preceding consonants is a phonological process of feature-spreading or a coarticulatory effect. In this respect, two sets of predictions emerge. The first is that a feature-spreading process should be quantitatively categorical. This means that assimilated consonants should be articulated with the same tongue configuration and with the same degree of posterior constriction as their underlyingly emphatic counterparts. If the assimilating emphatic consonant is produced with uvularization, the assimilated consonants should also be uvularized, and likewise, if it is produced with pharyngealization, the assimilated consonant should also be pharyngealized. Taking the extent of dorsal raising/retraction (in the case of uvularization) or tongue root retraction (in the case of pharyngealization) as an index of the degree of posterior constriction, the assimilated consonant should demonstrate the same magnitude of retraction in order to classify the process as feature-spreading. Furthermore, this categoricalness should be

consistent across plain consonants in emphatic context (e.g. C₁ in /C₁VC^s/) and phonemically emphatic consonants in both plain context (e.g. C₁ in /C^s₁VC/) and emphatic context (e.g. C₁ in /C^s₁VC^s/).

On the other hand, if the effect exerted by emphatics is coarticulatory, it should be gradient rather than categorical, with some variation between consonants belonging to the three categories just mentioned. This variation should be observed along two dimensions: the degree of constriction (i.e. the magnitude of articulator movement) and the location of constriction. In terms of its magnitude, the posterior articulator movement (be it the dorsum or tongue root) should be less extreme in coarticulated consonants than in phonemically emphatic ones in plain context, and it is likely that an emphatic consonant in emphatic context would be produced with either the same or greater posterior retraction. In terms of its location, the constriction in the coarticulated consonants should also be less posterior than the phonemic constriction in the triggering consonant.

3.3 Methodology

The ultrasound experiment is designed to address all research questions of this dissertation in a single data collection session per speaker. Hence, the methodology described in this chapter was employed in the two following chapters as well, but obviously using different words in each chapter. The wordlists of all research questions were merged into one randomized list.

The study in this chapter seeks to examine the tongue configuration in the production of plain and emphatic consonants in different phonetic environments. More precisely, it aims to assess the difference in dorsum and root position in coronals that contrast in emphasis /t s t^s s^s/ and to compare them to the same consonants in an emphatic context. This examination of the location and degree of lingual displacement is performed at different time points during the target consonants spoken by native speakers of Eastern Peninsular Arabic.

3.3.1 Participants. Data were collected from 15 speakers (5 females and 10 males) aged 19-52 ($\bar{x}$ = 28.6 years, SD= 9) at the time of recording. Participants all spent most of their lives in Arabic-speaking environments and 0.3-6 years in non-Arabic-speaking environments, mainly the English-dominant regions of North America. Their native dialect was Eastern Peninsular Arabic (sub-varieties: 13.3% (2/15) Bahraini, 26.6% (4/15) Al-Ahsa, 40% (6/15) Qatifi, and 20% (3/15) Dammami) and they resided for an average of 25 years in their native regions. In terms of their Arabic experience, although all participants were living in English-speaking regions at the time of recording, all reported speaking and listening to Arabic on a daily basis during the last 12 months. Nearly 60% of the speakers reported speaking other sub-varieties of Eastern Peninsular Arabic (which are not typologically different), and were exposed to other varieties of Arabic such as Egyptian (40% or 6/15 of participants); Levantine (53.3% or 8/15); Hijazi (33.3% or 5/15) and other less common dialects. This exposure can be through social interaction, media and other means. Spoken languages other than Arabic that were reported include English (100% of participants); Persian (20% or 3/15); French (13.3% or 2/15); and Italian, Spanish and Turkish (1 speaker each) with varying proficiency levels and ages of learning. None of these additional languages have pharyngeal or pharyngealized consonants.

3.3.2 Stimuli. The wordlist used in this chapter consisted of 22 disyllabic and trisyllabic words in which the target consonant occurs word-initially. This section presents the structure of the stimuli and the selection criteria that were implemented to avoid any confounds.

3.3.2.1 Stimuli conditions. The wordlist consists of four test conditions with the sequence C₁VC₂ where C₁ is the target consonant and C₂ is the source. The word-initial target coronal is either plain /t s/ (C₁) or emphatic /tˤ sˤ/ (C₁^ˤ) and the source consonant is either plain /t s b f/ (C₂)

or emphatic /tʰ sʰ/ (Cʰ₂). The consonants were cross-matched to constitute the following four testing conditions:

- A. Plain: Plain target in a plain environment, C₁VC₂
- B. Non-distinctive emphasis: Plain target in an emphatic environment, C₁VCʰ₂
- C. Distinctive emphasis: Phonemically emphatic target in a plain environment, Cʰ₁VC₂
- D. Double emphasis: Phonemically emphatic target in an emphatic environment, Cʰ₁VCʰ₂

The complete wordlist is presented in Table 3.1, in which words are grouped by test condition. Finding appropriate words meeting the double emphasis condition (containing the sequences /tʰVsʰ/ or /sʰVtʰ/) and the selection criteria (discussed in §3.2.2.2) was challenging because such words are rare in the language. There are many words with two phonetically emphatic consonants, e.g. [əsʰtʰaħab] ‘he/it accompanied’ and [əsʰtʰabar] ‘he/it is patient’, but in many of them, the second pharyngealized consonant is underlyingly a plain /t/ that surfaces as emphatic [tʰ] because it is immediately preceded by a phonemically emphatic consonant. This alternation will be further elaborated in Chapter 6 as it is related to the findings of the present study.

The emphatic consonants /ðʰ/ and /dʰ/ were excluded from this study because very few words contain them, and because their realization is dialectally and sociolinguistically variable in the Eastern Peninsular region (as both are realized as [ðʰ] by most speakers).

Table 3.1

The complete wordlist used in Experiment 1 showing the four test conditions (stress in boldface)

Condition	Target	Source	Example	Gloss	Source position
Plain	t	t	tat.bi:l	using spices	Coda
Non-distinctive emphasis	t	tʰ	tatʰ.bi:l	drumming	Coda
Distinctive emphasis	tʰ	b	tʰa.bi:b	doctor	Onset

Condition	Target	Source	Example	Gloss	Source position
Double emphasis	t ^ʃ	t ^ʃ	ba.t ^ʃ a.t ^ʃ es. ha:	her potatoes	Onset
Plain	t	t	tat.li: f	ruining	Coda
Non-distinctive emphasis	t	t ^ʃ	tat ^ʃ . fi:f	cheating in weight	Coda
Distinctive emphasis	t ^ʃ	f	t ^ʃ a. fi:f	shallow	Onset
Double emphasis	t ^ʃ	t ^ʃ	fəs.t ^ʃ at ^ʃ . ha:	her tent	Coda
Double emphasis	t ^ʃ	t ^ʃ	gi .t ^ʃ at ^ʃ	cats	Coda
Plain	t	s	ta.sa. fi:r	travelling-PL	Onset
Non-distinctive emphasis	t	s ^ʃ	ta.s ^ʃ a. fi:r	whistling-PL	Onset
Distinctive emphasis	t ^ʃ	s	t^ʃas	went away-MS	Onset
Double emphasis	t ^ʃ	s ^ʃ	(not found)		
Plain	t	s	tas. mi:m	poisoning	Coda
Non-distinctive emphasis	t	s ^ʃ	tas ^ʃ . mi:m	design	Coda
Distinctive emphasis	t ^ʃ	s	fa .t ^ʃ as	died- MSC	Coda
Double emphasis	t ^ʃ	s ^ʃ	(not found)		
Plain	s	t	sa.tar. ha:	covered her	Onset
Non-distinctive emphasis	s	t ^ʃ	sa.t ^ʃ ar. ha:	hit her on the face	Onset
Distinctive emphasis	s ^ʃ	f	s ^ʃ a. fi:r	whistle	Onset
Double emphasis	s ^ʃ	t ^ʃ	(not found)		
Plain	s	t	sat. ra:t	covering-FEM	Coda
Non-distinctive emphasis	s	t ^ʃ	sat ^ʃ . ra:t	hitting on face-FEM	Coda
Distinctive emphasis	s ^ʃ	m	s ^ʃ a. mi:m	core	Onset
Double emphasis	s ^ʃ	t ^ʃ	(not found)		
Plain	s	s	(not found)*		
Non-distinctive emphasis	s	s ^ʃ	(not found)*		
Distinctive emphasis	s ^ʃ	s	(not found)*		
Double emphasis	s ^ʃ	s ^ʃ	rə.s ^ʃ as ^ʃ . ha:	her bullets	Coda

* These sequences exist in the language but the words containing them do not conform to the stimuli selection criteria.

3.3.2.2 Stimuli selection criteria: Minimizing noise. Phonetic, phonological, prosodic and lexical factors were considered in selecting the target words. Nevertheless, some factors were better controlled than others because there is a trade-off relation between them.

3.3.2.2.1 Minimizing phonetic and phonological variation. In terms of phonetic and phonological factors, the distance between the source and target consonant is limited to one vowel only /a/. The choice of this vowel is based on evidence showing that it is more sensitive to coarticulation than /i u/ ([Jongman et al., 2011](#)), likely due to tongue body depression in its production which is compatible with articulatory demands of pharyngealization. The length of the intervening vowel is consistently short across the entire wordlist.

Also, none of the words contain a guttural consonant: uvular /q/, pharyngeal /ħ ʕ/ and laryngeal /ʔ/. The reason is that gutturals, according to ([McCarthy, 1991](#)), share similar acoustic and articulatory properties with pharyngealized consonants. Laryngeal /h/ had to be included because it is part of the suffix /-ha:/ ‘her’ that was necessary to control for stress in cases in which there was no alternative word or morphological template to control for it. The presence of the laryngeal /h/ is not expected to confound the results of the study because, unlike pharyngeal and emphatic consonants, the laryngeals do not involve any pharyngeal constriction ([Zawaydeh & de Jong, 2011](#)) and they do not induce coarticulatory acoustic effects on F1 and F2 of the preceding vowel ([Alwan, 1989](#); [Laufer & Baer, 1988](#)). Furthermore, in the few cases where an /h/ was included, it was always far away from the target consonant. For example, the target /s/ in /sat^ʕarha:/ ‘hit her on the face’ is distant from the laryngeal fricative and they are separated by an emphatic consonant.

3.3.2.2.2 *Minimizing prosodic variation.* The prosodic factors that are taken into consideration are stress, consonant position within the syllable and the syllabic distance between the target and source consonants.

In terms of stress, the target and source consonants always occur in unstressed syllables. For example, the target [t] and source [tʰ] in /tatʰ.bi:l/ ‘drumming’ are unstressed since the final syllable is heavy. Stress was manipulated either by suffixation (i.e. adding a suffix with a heavy syllable, e.g. /ha:/ ‘her’, to shift stress to a syllable other than the one containing source and target consonants) or by selecting a morphological template with a stress pattern abiding to the study criteria.

In addition, the positions of the target and source consonants in the syllable are controlled for. The target consonant is consistently at the syllable onset and the source can be at either syllable onset (e.g. /ta.sʰa.fi:r/ ‘whistling’) or syllable coda (e.g. /tasʰ.mi:m/ ‘design’). In order to avoid any correlation between syllable position of the source consonant and other factors that are more essential to the study (e.g. emphasis condition), the position of the source consonant was counter-balanced across all sub-conditions. Each emphasis condition contains words with the two source positions, for example, the non-distinctive emphasis condition consists of /ta.sʰa.fi:r/ and /tasʰ.mi:m/ where the source occupies the syllable onset and coda, respectively.

The third prosodic factor is the syllabic distance between the two consonants under investigation, which is correlated with the previous factor, syllable position. Source and target consonants can be either in the same syllable (e.g. /tasʰ.mi:m/ where /t/ and /sʰ/ are the target and source, respectively) or in two different syllables (e.g. /ta.sʰa.fi:r / with target /t/ and source /sʰ/). The two factors (i.e. consonant position and syllable distance) are correlated: whenever the source consonant is in the coda, it co-exists with the target in the same syllable. The opposite is also true:

whenever the source consonant is in the onset, the target is in a different syllable. It is not feasible to keep the source position and syllabic distance constant across all words without violating other factors. However, syllabic distance is not expected to affect the results because evidence in the literature shows that the assimilation of emphasis crosses syllable boundaries, and even word boundaries in some dialects (e.g. [S. Davis, 1993](#); [S. Davis, 1995](#); [Younes, 1994](#)).

3.3.2.2.3 Minimizing lexical variation. Lexical factors such as word frequency and neighborhood density were not controlled, for various reasons. Since it is already difficult to come up with a balanced wordlist that abides by all phonetic and phonological factors, it is not feasible to simultaneously control for lexical factors. Another problem is the trade-off relation between these factors and other fundamental phonetic and prosodic factors. For example, selecting words with a particular frequency may require including words with varying stress patterns, different vowel lengths, etc. Moreover, there is no tool for measuring frequency and neighborhood density in the regional dialects of Arabic, including the one under investigation. Lexical databases designed for Modern Standard Arabic such as Aralex ([Boudelaa & Marslen-Wilson, 2010](#)) cannot be used for spoken regional dialects.

3.3.3 Data collection. Data collection took place at two locations but using the same instrumentation and setup: the Sound Patterns Laboratory at the University of Ottawa and the Phonetics Laboratory at the University of Toronto. Prior to the start of the ultrasound recording, participants were trained on the test words in order to familiarize them with their orthographic form and to reduce the effect of word frequency. The familiarization phase involved presenting a written definition, synonym or antonym of the target word on a computer screen along with two words choices. Participants were asked to read the definition and select the correct word by pronouncing it.

Mid-sagittal ultrasound images were generated using Terason t3000 machine operating Ultraspeech 1.2 ([Hueber, Chollet, Denby, & Stone, 2008](#)) in a direct-to-disk mode. Although the more recent version 1.3 of Ultraspeech includes many improvements such as a higher ultrasound image resolution, it has incompatibility issue issues with our Terason t3000 acquired in 2009 (not reported in other labs using a more recent version of Terason t3000): whether using Direct-to-Disk or RAM-to-Disk mode, the output audio file contains no acoustic signal allowing subsequent synchronization. To overcome this problem, I used the older version of Ultraspeech (1.2), which records the first 14 seconds of the session, and I simultaneously recorded the entire session on another audio channel via an external PC. The Ultraspeech audio was recorded using a Shure Beta 54 head-mounted condenser microphone and the full-session audio was recorded using a Shure Beta 53 head-mounted condenser microphone into a PC running Audacity. Both microphones were mounted on the headset and each one was connected to a Shure X2u preamplifier/digital-to-analog converter/USB interface. Audio signals were recorded with a 16-bit depth and a 44.1 kHz sample rate. Using a micro-convex array transducer (8MC4 4-8 MHz) which was stabilized using an Articulate Instruments headset ([Scobbie, Wrench, & van der Linden, 2008](#)), the 320×340 pixel .bmp images were generated at 60 frames/second. The pressure exerted by the ultrasound transducer on the chin was kept as minimal as possible in order to ensure that participants could speak naturally.

The ultrasound recording session took place in a sound-attenuated booth where participants read a randomized list of words embedded in the carrier sentence /ʔana ʔagu:1... əlfahad/ ‘I say... to Fahad’. The stimuli were presented on a computer screen and each sentence appeared for 3.5 seconds (in order to make speech rate as normal as possible) and was repeated 5 times with a short break between every two repetitions. The recording session took about 24 minutes in addition to

self-paced breaks between each two blocks of words, producing 7-9 GB of ultrasound images per session for all research questions in this dissertation. Prior to displaying the stimuli, participants' occlusal plane images were acquired by placing a tongue depressor between the upper and lower teeth and asking the participant to press their tongue against it while biting down.

3.3.4 Data processing. The first step in processing the data was to align the two audio recordings since the external audio started a few seconds earlier than Ultraspeech audio. The purpose of this audio alignment was to remove the extra time at the start of the external audio in order to ensure that the time stamps of the ultrasound frames corresponded to times in the external audio recording. As illustrated in Figure 3.1, the same acoustic event, such as the release of an abrupt /t/, was located on the two audio recordings and the difference between these two time points was the amount of time that was then removed from the external audio recording, which was used in all subsequent data processing procedures.

Semi-automatic segmentation of the aligned audio was then conducted using the Montreal Forced Aligner ([McAuliffe et al., 2017](#)). Since the aligner did not include a pretrained model for Arabic, a Romanized grapheme-to-morpheme model was used to construct a pronunciation dictionary which was then used for auto-segmenting the words in the audio data. Phone interval boundaries were corrected manually in Praat ([Boersma & Weenink, 2018](#)). A Python script ([Mielke, 2015](#)) used the resulting textgrid, with reference to the time stamp in the original file name, picked all ultrasound frames that occurred within labeled phone intervals in the textgrid, enlarged them by 150%, renamed the files and converted the .bmp images to .jpg, thus, making them compatible with Palatoglossatron software ([Baker, 2005](#)).

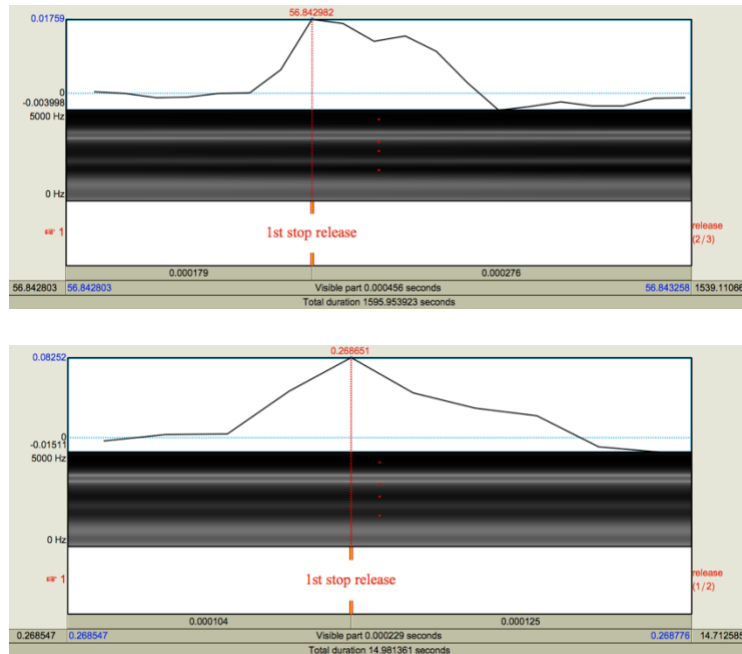


Figure 3.1. Alignment of audio recordings by locating the stop release on the external audio (top) and on Ultraspeech audio (bottom). An interval of length 56.842982 (top) minus 0.268651 (bottom) was removed from the start of the external audio.

In order to select ultrasound frames corresponding to specific acoustic events, a Praat script was used to mark the release of all stop consonants (this was then corrected manually), and to replicate the interval boundaries of all other consonants in a new tier. This was done for both target and source consonants. A Python script used this 3-tier textgrid to select the following frames of interest. For all stop consonants, the script selected the closest frame to the stop release (henceforth, 100%) and another frame at (or near) mid-closure (50%). Three frames were selected from all other consonants – including fricatives, nasals, liquids – and the intervening vowel: at 25, 50 and 75% of the segment duration (with 0 being the onset and 100 the offset). Results of vowel frames are not reported in this study.

With this frame selection, the total size of the data in this dissertation was 29,331 ultrasound images: 14,198 images for the study in Chapter 3 and 4 together (with some images used for both), and 15,133 in Chapter 5. The size of the data analyzed in this chapter specifically was 7,950 images: 2,100 for the plain condition; 2,100 for non-distinctive emphasis; 2,250 for distinctive emphasis; and 1,500 for double emphasis. Tongue contours in all selected frames were traced manually using Palatoglossatron, and X and Y coordinates the points in each tongue trace were extracted for further analysis.

3.3.5 Data analysis. Data extracted from Palatoglossatron were analyzed with Smoothing-Spline analysis of variance ([henceforth, SS-ANOVA; Gu, 2002](#)), which was first applied by [Davidson \(2006\)](#) to ultrasound data. This was used in this dissertation to assess statistical differences between groups of tongue contours. The groups contours were considered significantly different if Bayesian confidence intervals did not overlap at the portion of the tongue spline in question; for example, tongue splines can be significantly different at the tongue body but not at the tongue root if their confidence intervals overlap there. Prior to conducting SS-ANOVA, 1) tongue splines were rotated to make the occlusal plane horizontal and 2) the spline coordinates were converted from a Cartesian to a polar coordinate system. These two procedures are explained in more details below.

The occlusal plane angle was measured as the number of degrees by which the occlusal plane is rotated forward – that is, the angle between highest point along the occlusal plane (on the right of the image in Figure 3.2) and the lowest point to left tongue curve. Angles were measured for each individual speaker independently and ranged between 0° and 25°.

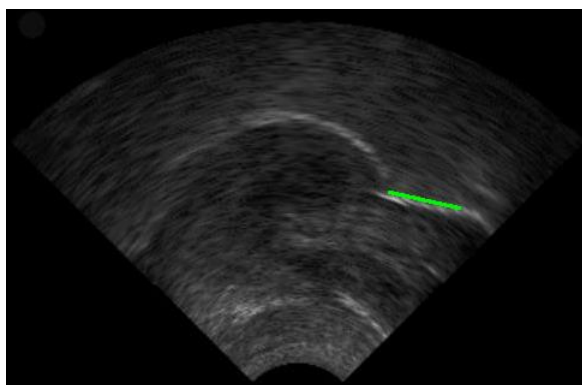


Figure 3.2. Ultrasound image illustrating the measurement of the occlusal plane angle per speaker as indicated by the green line.

These angles were used to rotate the tongue splines of the respective speaker so that splines are aligned across speakers, given that the transducer was titled backward to capture the tongue root area. An example of rotated (by 14.25°) and un-rotated tongue splines is provided in Figure 3.3 where the x-axis represents tongue anteriority and y-axis tongue height. The rotated splines were used in all subsequent analyses. In cases where the ultrasound transducer was adjusted during the recording session, I imaged the occlusal plane for the new transducer position. In such cases, the speaker had more than one angle, each corresponding to a specific phase of the recording and used to rotate the images of the respective phase. For example, if the occlusal plane angle in the first round of the recording was 11° and, then adjusted to 13° , the ultrasound images recorded in the first round were rotated to 11° and the rest of images to 13° . In such cases, the images of the same consonant that was produced before and after the transducer re-adjustment were compared to each other in terms of tongue height and anteriority (or blade position) and if the shadow of hyoid bone or mandible appears after the re-adjustment. This comparison was to verify if there was any change along any of these dimensions that requires removing either portion of the data.

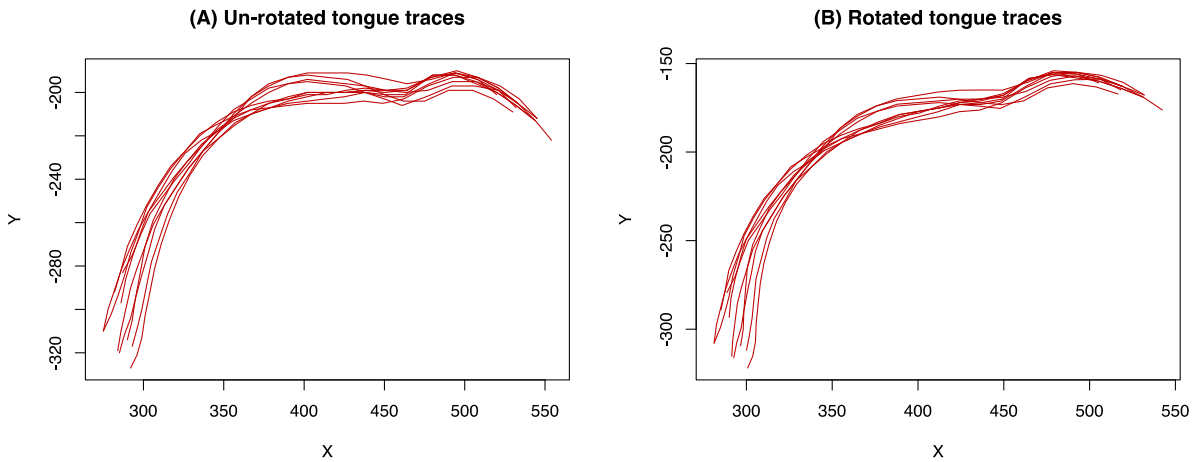


Figure 3.3. Original tongue traces from speaker S06 showing (A) original un-rotated splines and (B) splines rotated to 14.25° .

3.3.5.1 Using polar versus Cartesian SS-ANOVA. Following [Mielke \(2015\)](#), SS-ANOVAs were computed using polar coordinates and the plots were generated using Cartesian coordinates. Ideally, assessing differences of tongue splines involves comparing each part of the tongue according to the direction of its movement. For instance, the tongue body raising involves a change of the spline along the vertical axis where a low tongue body has a smaller y value than a high tongue body; likewise, tongue root retraction involves a change along the horizontal dimension where a retracted root has a smaller x value than an advanced root. According to Mielke, doing this type of comparison with Cartesian coordinates at the tongue root is not adequate because the spline along this area is oblique to the x-axis (Figure 3.4 left) which does not reflect the direction of root retraction, unlike the tongue body where the spline is parallel to the x-axis. Instead, in polar SS-ANOVA (Figure 3.4 right), the tongue root is perpendicular to the semi-horizontal radius, which mimics the direction of the root movement (advancement and retraction). Because of this inaccuracy of Cartesian SS-ANOVA comparison at the tongue root and because this study is interested in this particular portion of the tongue, polar SS-ANOVA is deemed more

appropriate for the present analysis. The overall tongue spline using polar coordinates is similar to the actual tongue traces from which the plots in Figure 3.4 are derived.

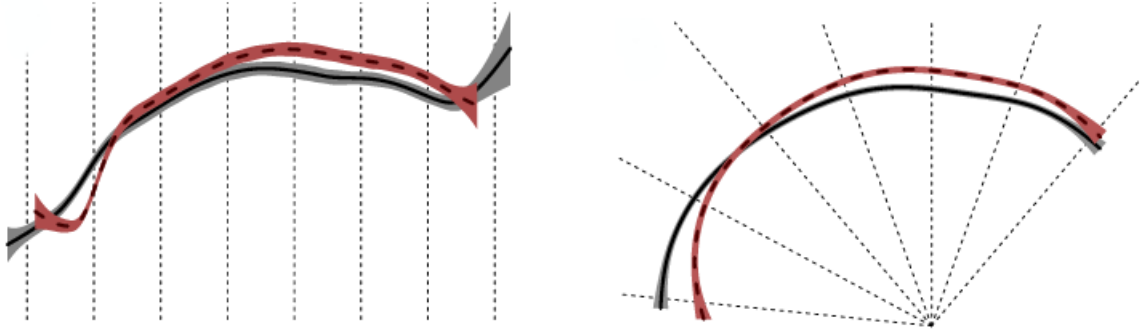


Figure 3.4. SS-ANOVA using Cartesian (left) and polar coordinates (right), reproduced from [Mielke \(2015\)](#).

In polar SS-ANOVA, any point along the tongue contour is expressed in two polar coordinates: the radius r which refers to the distance from a reference point and the angle θ between the radius and the polar axis. The reference point, called ‘origin’ in this dissertation, was determined as the point on the plane whose x and y coordinates corresponded to the mean of all x values included in the SS-ANOVA comparison and the minimum of all y values, respectively. After exporting the Cartesian x and y from Palatoglossatron for all the traces used in this study and rotating them to the occlusal plane angle, they were converted to r and θ using the trigonometric function, tangent, as follows:

$$r = \sqrt{(x^2 + y^2)}$$

$$\theta = \tan^{-1} \frac{y}{x}$$

Polar r and θ were converted back to Cartesian x and y for plotting SS-ANOVA results using the sine and cosine functions: $x = r \times \cos(\theta)$ and $y = r \times \sin(\theta)$, respectively.

3.3.5.2 Limitations of SS-ANOVA. There are a few limitations to the application of SS-ANOVA on this type of data. One of these limitations is that, unlike some other statistical tests

such as regression models, it does not allow testing the interaction between independent variables. For instance, a two-way interaction between the phonetic context (with two levels: plain vs. pharyngealized context) and the identity of the target consonant (e.g. [b] vs. [f]) – as they affect the tongue shape – is not possible in SS-ANOVA. To overcome this, these interactions were generated manually by creating a third variable containing the following values: [b] in plain context, [b] in pharyngealized context, [f] in plain context and [f] in pharyngealized context. The new variable is used in SS-ANOVA as its only testing variable. Although this apparently mimics the interactions used in mixed-effect models, for example, it may not yield the same significance testing results. Another issue with SS-ANOVA (and other statistical techniques applied to this specific type of ultrasound data) is that it allows only within-speaker comparisons because the tongue contours of different speakers cannot be normalized.

A further limitation relates to the confidence intervals which, in most (or all) SS-ANOVA tests conducted here, are very small, hence, increasing the chance of obtaining a significant difference between the portion of the tongue splines under investigation. Two reasons may have contributed to the small confidence intervals in this study. First, the size of the data was large enough to pick very small differences in tongue splines. The second is the use of the polar system to conduct the test; this is evident when comparing the confidence intervals in the Cartesian and polar SS-ANOVAs provided in Figure 3.4. According to Mielke (email communication), the confidence intervals in Cartesian SS-ANOVA is larger than in polar SS-ANOVA because in the Cartesian variant, the part of the tongue that approaches a vertical position has inflated standard errors. This is because, when measuring the horizontal differences along the tongue root as in Cartesian-based analysis, there is normally a large variation at any particular point along the x-

axis value. This means that the part of the tongue that is vertical has a large variation along the y-axis inducing greater standard errors and, consequently, larger confidence intervals.

Two tongue splines were considered to be different from each other at a given portion of the tongue if their confidence intervals do not overlap and the size of the difference was determined according to the distance between the tongue fits.

3.3.5.3 Additional analyses. Given the limitations to SS-ANOVA, additional analyses were performed to supplement SS-ANOVA results. These include: 1) quantifying differences based on SS-ANOVA fits (presented in scatterplots); 2) a relative measure of tongue root retraction based on plain C and emphatic C^s reference points; and 3) performing ANOVA on the radius at the tongue root area to compare testing conditions (presented in boxplots along with the ratios).

3.3.5.3.1 Scatterplots for quantifying differences of tongue fits. The purpose of the scatterplots was twofold: to quantify the differences between testing conditions by using the SS-ANOVA tongue fits at a particular portion of the tongue and to summarize these differences across all speakers instead of presenting individual speakers' SS-ANOVA plots. To illustrate how they are generated and what they represent, consider the sample SS-ANOVA plot in Figure 3.5 and the corresponding scatterplot in Figure 3.6. The SS-ANOVA plot presents the tongue splines by one speaker for three example conditions: Plain, Coarticulated and Emphatic. The tongue fit at the root area for the Coarticulated condition is more retracted (i.e. lower x value) than the Plain condition and less than the Emphatic condition (i.e. higher x value).

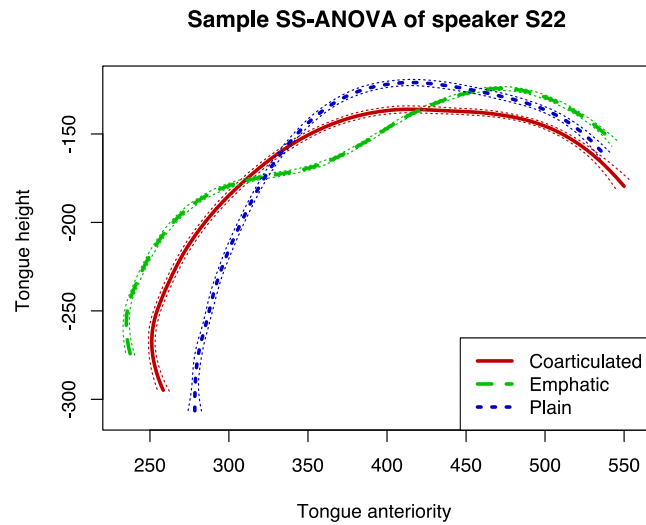


Figure 3.5. SS-ANOVA fit (tongue tip to the right) and 95% CI of three example conditions.

Within-speaker comparisons of the three conditions shown in Figure 3.5 are summarized in the scatterplot in Figure 3.6 where the points represent individual speakers. Among the different tongue portions, the tongue root and body were chosen. Each panel of Figure 3.6 represents the difference between two of the three conditions. Following [Mielke, Carignan, and Thomas \(2017\)](#), the x-axis represents the minimum difference (per speaker) between SS-ANOVA fits for the two respective conditions (e.g. fit of Emphatic condition minus fit of Coarticulated condition) at an angle from horizontal to 45° ($0, \pi/4$) measured at the polar origin of the corresponding SS-ANOVA fit. This angle was taken to represent the tongue root area, and a difference between two SS-ANOVA fits at this area represents a mostly horizontal tongue root displacement (i.e. retraction/advancement). Similarly, the y-axis of the scatterplot refers to the minimum difference (per speaker) between the two conditions at the tongue body area, the angle from 75° to 105° ($\pi/2.4, \pi/1.7$). A difference between two SS-ANOVA fits at this area represents a change in tongue body height. Defining the angles of the tongue root and body was based on a visual inspection of

the tongue contours of all speakers where the articulator in question falls within these angles. These angles apply to all scatterplots in this dissertation except those in §3.4.1 (about uvularization and pharyngealization) where the comparison is conducted at the tongue dorsum 45° to 75° ($\pi/4$, $\pi/2$).⁴ instead of the tongue body. The dorsum was chosen specifically in that section because it presents the distinction between pharyngealization (at the tongue root) and uvularization (at the dorsum). The reason for choosing the tongue body in most scatterplots about assimilation was that, according to individual speakers' SS-ANOVA plots, it represented another articulatory correlate of emphasis which seemed to accompany the tongue root.

For interpreting all scatterplots, the horizontal and vertical dashed lines represent zero difference at the tongue body and root, respectively. In the scatterplot of Plain minus Coarticulated (right panel of Figure 3.6 and similar plots), a positive difference along the x-axis means that the Coarticulated condition has a more retracted root than the Plain condition, and a negative difference means that the Coarticulated condition has a less retracted root than the Plain condition. In the scatterplot of Emphatic minus Coarticulated (left panel of Figure 3.6 and similar plots), a negative difference along the x-axis means that the tongue root position in the Coarticulated condition is less retracted than the Emphatic condition, and vice versa.⁴

⁴ Figuring out gradient/categorical effects in the scatterplot may not be intuitive. 1) If the tongue root in the affected consonant is different from *both* the Plain and the Emphatic condition, the effect is considered gradient. 2) If the tongue root in the Coarticulated condition is different from the Plain consonant and similar to the Emphatic one, it is considered as categorical assimilation. 3) If the tongue root in the Coarticulated condition is similar to the Plain condition, it is considered as resistance to assimilation.

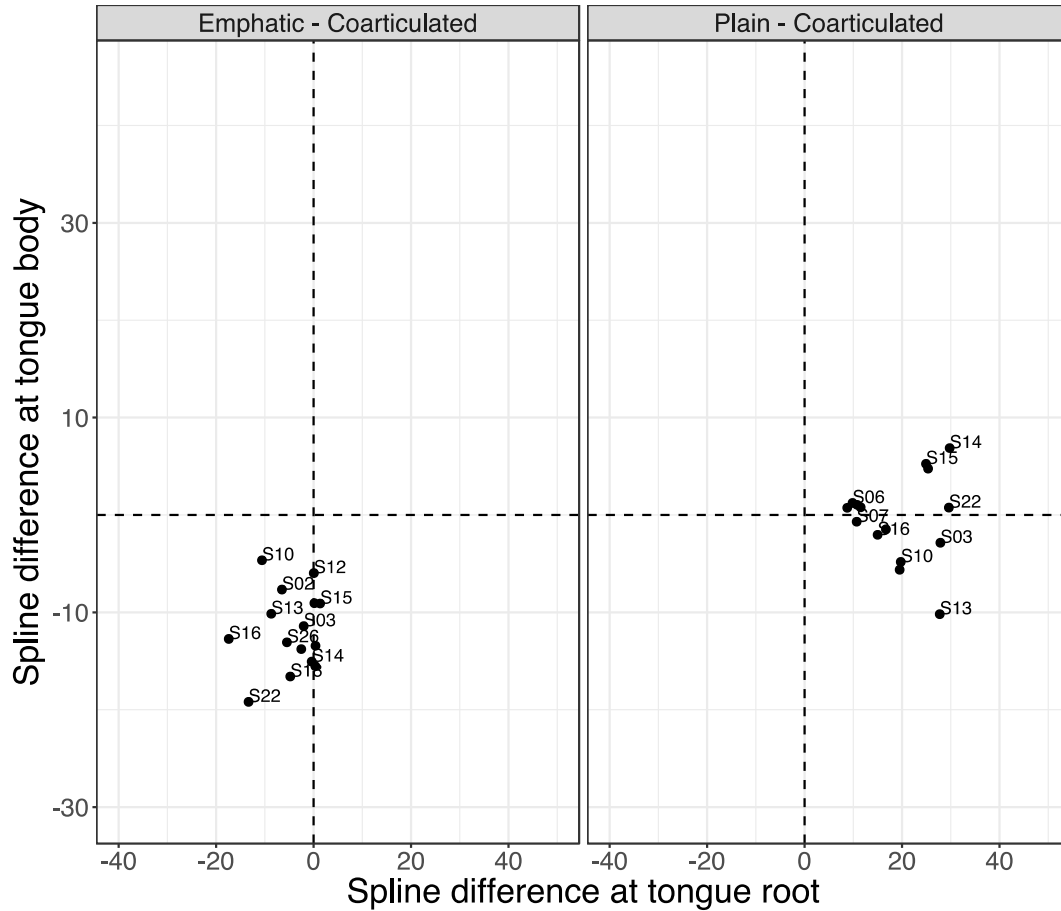


Figure 3.6. The difference between SS-ANOVA fits of the three conditions shown in the sample SS-ANOVA for all speakers. A negative root difference in the left means Coarticulated is less retracted than Emphatic and a positive root difference in the right means Coarticulated is more retracted than Plain.

3.3.5.3.2 *Retraction ratios for the size of coarticulation relative to plain and emphatic references.* The purpose of using a tongue root retraction ratio was to estimate the magnitude of coarticulation relative to two reference/extreme points: the plain condition (represented with 0) and the emphatic one (represented with 1). Consonants with no coarticulation should have a root retraction ratio close to 0, and consonants with 100% coarticulation should have a root retraction ratio close to 1. I present the root retraction ratio results in boxplots throughout this dissertation, and the boxplots of all consonants by all speakers can be found in the Appendices.

For every speaker, a polar origin was determined as the minimum y-value and mean x-value and, then, the trace was converted into polar coordinates. The position of the tongue root (which is the principal articulator in emphatics as shown later) corresponds to the radius from the polar origin to the tongue trace at the angle between 0° and 45° (0, $\pi/4$). For every tongue trace, I obtained the average radius of all radii within the tongue root region and calculated the z-score of this average to normalize the tongue size per speaker:

$$zR = (R - \text{mean } R) / \text{mean } R$$

where R refers to the average radius of a given trace. This normalized average radius (zR) was then used to calculate the ratio with the following formula, where the ‘mean zR of plain’ and ‘mean zR of emphatic’ refer to the mean normalized radius of all traces in the plain and in the emphatic condition, respectively.

$$\text{Retraction Ratio} = \frac{\text{zR of a given token} - \text{mean zR of plain}}{\text{mean zR of emphatic} - \text{mean zR of plain}}$$

3.3.5.3.3 ANOVA for categorizing coarticulation and assimilation. The purpose of performing within-speaker ANOVA was to statistically categorize the assimilatory condition as either resistant, gradient or categorical. The dependent variable was the mean radius of all tokens of a given condition at the tongue root region, which indicates the amount of tongue root forward or backward displacement.

The ANOVA test for the difference between a given target consonant in an emphatic context (assimilation condition) and in plain context was used to assess whether the consonant is resistant or not. If the difference between the assimilation and plain condition is small ($F < 30$), the consonant is considered resistant and if the difference is greater ($F > 30$), the consonant is considered assimilated/coarticulated.

Another ANOVA test was performed to compare the assimilated target consonant to the distinctively emphatic trigger and to assess whether the effect is gradient or categorical. If the difference between the assimilated consonant and the emphatic trigger is small ($F < 30$), the effect is considered categorical, and if the difference is greater ($F > 30$), the effect is categorized as gradient.

In terms of the criterion for determining whether a given ANOVA difference is small or large, it appears that the p-value by itself did not yield enough information for categorizing the assimilatory effects unlike F-distribution which varied considerably across consonants; therefore, I used F-value as a criterion for categorizing the size of the effect. Given the large sample size, the p-value almost always indicated a significant result at $\alpha=0.001$ even if the difference between the consonant categories was trivial. For example, the assimilated /s/ was produced with a substantial tongue root retraction similar to that of distinctive /s^h/ by speaker S06 (as shown in Figure 3.7); however, the p-value suggested that the assimilated /s/ and the trigger /s^h/ were significantly different from each other at the tongue root: $F(1, 27)= 18.97, p < 0.0001$. The p-value in this case (and in some other cases) captured a difference that is real, but probably not linguistically meaningful. There were, also, examples in which the p-value would suggest the grouping of speakers who seemed to assimilate /s/ with speakers who produced a resistant /s/ into one category. To illustrate this, the ANOVA test of the difference between assimilated and plain /s/ by S12 speaker was $F(1,18)=14, p=0.001$ and by speaker S14 was $F(1,18)=138.8, p<0.001$. The p-value for both speakers suggested a significant difference between the /s/ categories although they had completely different F-values (14 vs. 138.8). Hence, relying on p-value for the ANOVA between assimilated and plain condition would always suggest a significant assimilatory effect even if the effect were very small. The F-value, on the other hand, provided a more straightforward indication

of the variance between the consonant categories in question (relative to the within-category variance) with a larger between-category variance yielding higher F-value. When examining the F-values of all ANOVAs, it appears that the F-values varied substantially across speakers and across consonant categories, which renders F-value more useful in categorizing the differences between consonants across speakers.

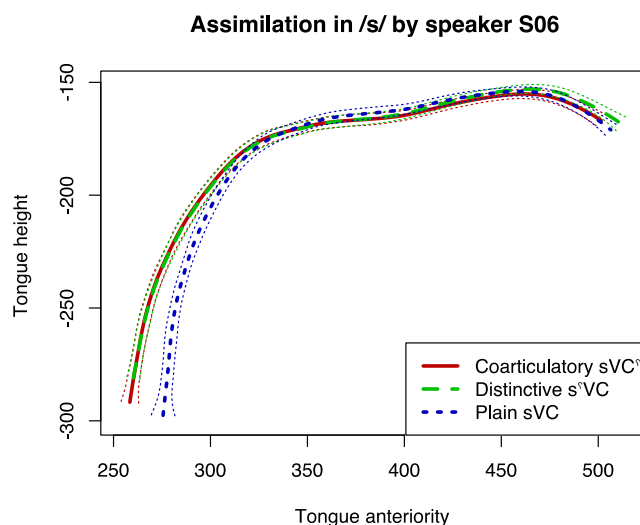


Figure 3.7. Example showing a significant, yet inconspicuous, difference between coarticulated and distinctive /s/'s.

The use of $F=30$ as a threshold for categorizing differences raises two questions: 1) why not consider consonant categories as similar when they are completely identical, i.e. $F=0$?; and 2) why use $F=30$ specifically? I address the first issue here based on linguistic grounds and the second based on statistical grounds.

Firstly, two consonant categories (e.g. assimilated C and distinctive C^s) were rendered similar not only if there was no difference at all between them (i.e. $F=0$), but also if the difference between them was small ($F < 30$). The decision to consider categories with a small statistical difference as linguistically similar is based on the linguistic phenomenon of incomplete

neutralization. This is a widespread phenomenon that has been attested in several languages, for example, final obstruent devoicing in German ([Nícenboim, Roettger, & Vasishth, 2018](#); [Roettger, Winter, Grawunder, Kirby, & Grice, 2014](#)) and in Russian ([Kharlamov, 2012](#)), English flapping ([Braver, 2014](#)), and Arabic palatalization ([Zellou, 2013](#)). What this means is that although two phonemically distinct phonemes merge phonologically and become one phoneme (with the phonological distinction being reduced), they tend not to be phonetically identical and some residue of their underlying forms remain on the surface realization. Take as an example the case of sibilant palatalization in Moroccan Arabic, e.g. /seʒera/ ~ /ʃeʒera/ ‘tree’, where the palatalized sibilant is produced with a slightly more forward place of articulation (acoustically, with a higher centre of gravity) than phonemically palatal sibilants. Therefore, it is possible that emphasis spread is a case of incomplete neutralization. If it is, we would expect an assimilated consonant not completely identical to distinctive emphatics, but maintain some distinction from distinctive emphatics at the (articulatory) phonetic level. This can be manifested articulatorily by producing an assimilated consonant with lesser backward retraction of the tongue root than distinctive emphatics, in which case the difference between their root radius should yield $F \geq 0$.

This leads to the second issue of what can be considered as a small difference between consonant categories. The decision to use $F=30$ as the cutoff point to classify differences between consonants was based on the overall F -values of all ANOVAs used in this study. $F=30$ corresponded to the rounded average first quartile of all ANOVA tests on the four consonants examined in this study (the unrounded value was 27-28). It yielded a classification of consonant conditions that is linguistically reasonable: based on $F=30$, consonants with similar tongue traces (by inspecting overlapping raw traces rather than SS-ANOVA fits) can be categorized as similar, and those with distinct traces can be deemed different.

Other indicators based on F-values (minimum F-value, mean F-value and mean F-value minus 2 standard deviations) could not be used as thresholds because they included too many or too few tokens. Using the minimum F-value (which was 0) assumes that assimilation necessarily leads to complete neutralization. On the contrary, using the mean F-value of all ANOVAs meant that even consonant categories with large F-values would be classified as similar. The same issue applied to M-2SD: using a threshold of 1-2 SD below the mean F-value meant that consonant categories with fairly large F-values would be categorized as similar. Since M-2SD is -307.5 and M-1SD is -100.9 which are relatively large, the assimilated consonants with a tongue shape dramatically different from distinctive C^s would be categorized as fully assimilated.

The threshold must be the same for all consonants used in this study (i.e. /t s l n/) and it is not plausible to use a different threshold for every consonant: because I compared the consonants with each other, using different thresholds can bias the categorization of consonant differences. That is, if the threshold of /s/, for example, is lower than that of /t/, the former consonant would tend to be categorized as categorical more often than /t/.

3.4 Results

The presentation of the statistical findings will start with the most fundamental articulatory characterization of emphatics – i.e. whether they are produced with uvularization or pharyngealization. This is followed by the results on the categoricalness or gradience of coarticulation. Given the aforementioned limitation of SS-ANOVA, the results will address each variable separately and for each participant individually. Hence, for each comparison, only one speaker's SS-ANOVA plot is shown provided that this plot is representative of the rest of the sample in terms of the tongue shape and the direction of the difference between the two testing

conditions under investigation. Individual speaker's SS-ANOVA plots are in Appendices 5-13. The main emphasis is on the tongue shape at mid-stop closure and mid-consonant for non-stops, although tongue shapes at other time points are also presented.

3.4.1 Uvularization vs. pharyngealization (Question 1). To answer this question, I will look at phonemically emphatic consonants in both plain and emphatic contexts, but not at plain consonants assimilated with an emphatic consonant. I will distinguish the contrastive (i.e. C^ʕ in C^ʕVC) and doubly emphatic (i.e. C₁^ʕ in C₁^ʕVC₂^ʕ) conditions from the plain condition to assess whether C^ʕs are produced with the tongue root or the dorsum.

Ultrasound data showed that emphatic consonants were articulated with a backward retraction of the tongue root towards the posterior pharyngeal wall (Figure 3.8 left). This was accompanied by a backward and downward displacement of the dorsum and a depression and flattening of the tongue body. This configuration is in accordance with the pharyngealization description of emphatics rather than with uvularization because the dorsum was not simultaneously raised and retracted. In contrast, the tongue root was in a neutral position and the convex tongue body was higher in the oral cavity while producing their plain counterparts /t s/ (Figure 3.8 right).

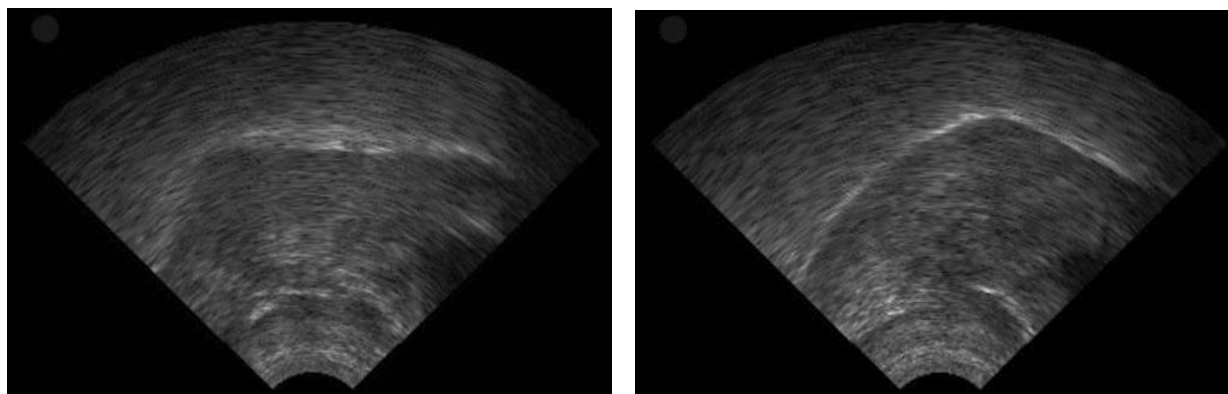


Figure 3.8. Ultrasound images of speaker S17 showing the tongue shape at the release of /tʕ/ in /tʕa.fi:f/ 'shallow' (left) and /t/ in /tatli:f/ 'ruining' (right).

SS-ANOVA at mid-phone/closure showed a substantial and statistically significant root retraction and tongue body and dorsum lowering in phonemically emphatic consonants /tˤ sˤ/, whether in plain or in emphatic context and compared to their plain counterparts /t s/ (Figure 3.9). This is true for all speakers with varying degrees of tongue root retraction and tongue body lowering. The tongue configuration in Cˤ is exactly the same at all time points of the consonant duration (i.e. at mid-closure and release of /tˤ/ and at 25, 50 and 75% of /sˤ/) for all speakers, suggesting a consistent degree of pharyngealization throughout the consonant (Figure 3.10). The degree of phonemic pharyngealization was the same for all speakers regardless of the phonetic context: tongue splines of distinctive Cˤ in plain context and in emphatic context (double emphasis) were completely overlapping.

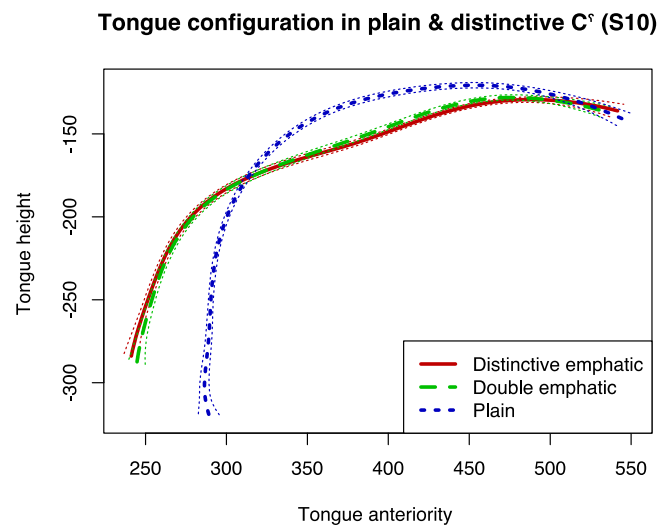


Figure 3.9. SS-ANOVA fit at mid-phone/closure (tongue tip to the right) and 95% CI showing Cˤ in plain and in emphatic context and plain C.

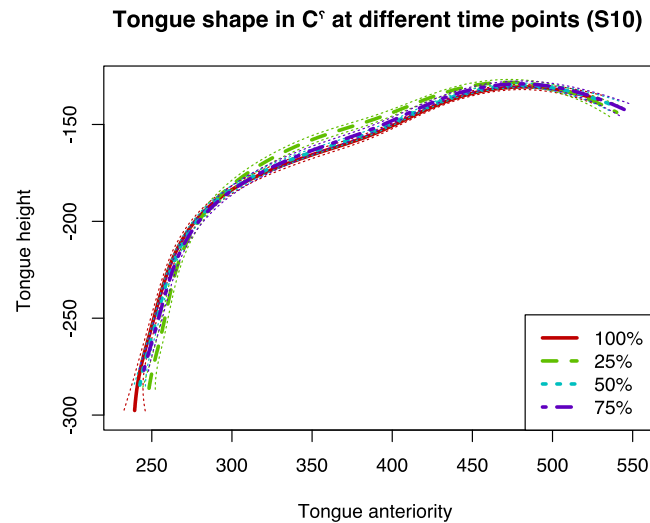


Figure 3.10. SS-ANOVA fit and 95% CI showing distinctive C^ɕ at all selected time points.

Within-speaker comparisons of the three conditions (as shown in Figure 3.9) are summarized in the scatterplot in Figure 3.11 where the points represent individual speakers. What the scatterplot in Figure 3.11 (right) illustrates is that, consistently across all speakers, there was a backward displacement of the tongue root in articulating distinctive C^ɕs compared to their plain counterparts. In other words, the direction of the difference between C^ɕ and C was the same for all speakers because their values are all below 0, and the negative values here indicate that the distinctive C^ɕ had a more retracted root than plain C (which has a greater value on the x-axis in the SS-ANOVA plot). They varied, though, in the magnitude of the tongue root displacement in producing C^ɕ which was well above 15 mm but did not exceed 60 mm. The tongue dorsum was also affected by emphasis. The dorsum was lower in distinctive C^ɕ than plain C for all speakers. This lowering, which did not exceed 5mm and which goes into the opposite direction of what studies supporting uvularization have found, is one of the articulatory correlates of producing pharyngealization. This supports the view that emphatics are produced by retracting the tongue

root and, therefore, they are pharyngealized rather than uvularized. The scatterplot in Figure 3.11 (left) indicates that all speakers produce the distinctive C^s in plain and in emphatic context with the same extent of tongue root and dorsum displacement. Now that it is determined that emphatic consonants are pharyngealized and to accurately describe these consonants in light of this finding, the term ‘pharyngealized consonants’ will be used in this dissertation instead of ‘emphatics’.

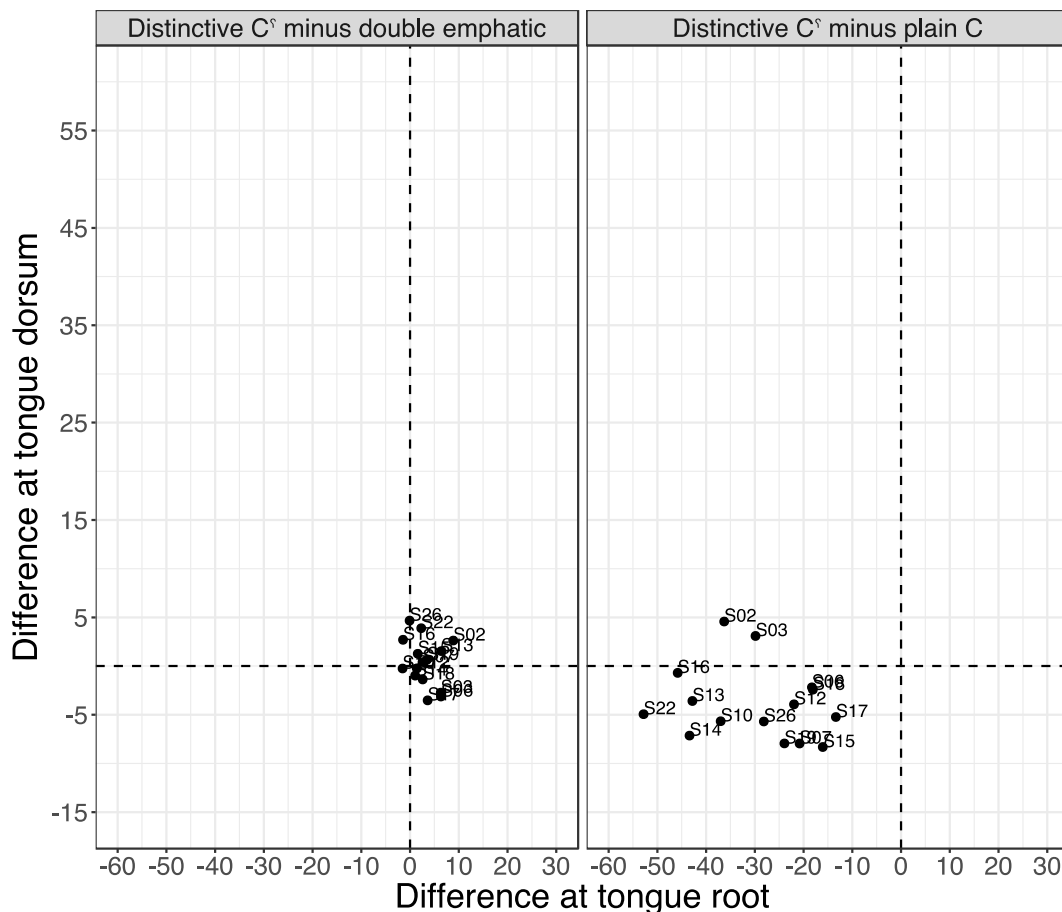


Figure 3.11. The difference between SS-ANOVA fits at mid-phone/closure for all speakers: distinctive C^s in plain context relative to the same consonant in pharyngealized context (left) and to plain C (right) where the negative root difference means more retraction in distinctive C^s .

In order to examine whether the extent of pharyngealization is influenced by the manner of articulation of the respective consonant, the two distinctively pharyngealized /t^s/ and /s^s/ were

compared to each other. Results revealed no significant difference between /t^s/ and /s^s/ at the tongue root and little or no difference at the tongue body area (Figure 3.12). As illustrated in Figure 3.13, this was consistent across all speakers except one (S03) who demonstrated a moderate difference between /t^s/ and /s^s/ at the tongue root (10-20mm). The plain /t/ was produced with a more forward tongue root than /s/ for about 73.3% of speakers.

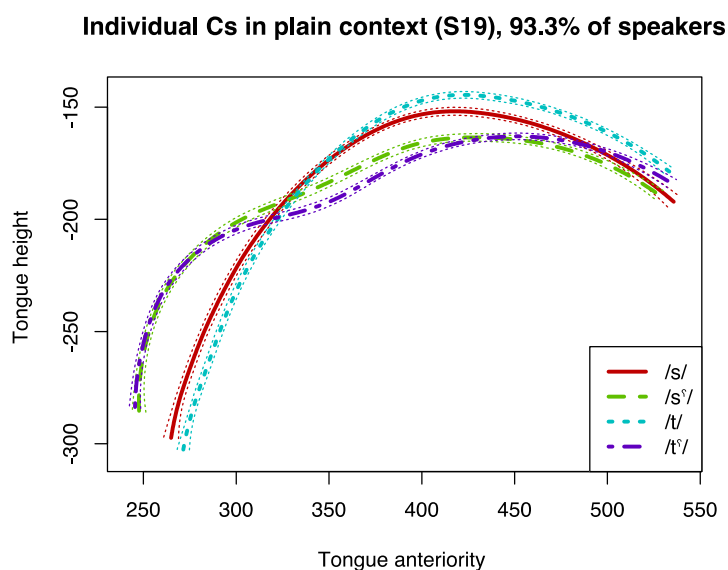


Figure 3.12. SS-ANOVA fit at 50% of the phone/closure and 95% CI showing /t s t^s s^s/ in plain context.

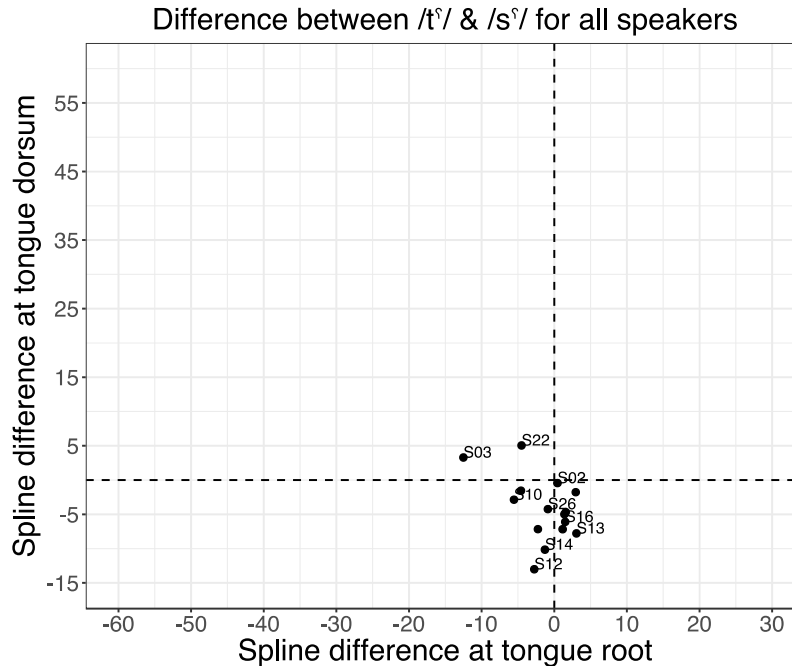


Figure 3.13. The difference between SS-ANOVA fits of /tʰ/ minus /sʰ/ in plain context for all speakers.

Positive root difference and negative body difference mean that the root in /sʰ/ is more retracted and the body is higher than in /tʰ/.

Also, comparison of pharyngealized C1 in CʰVC and C2 in CVCʰ revealed no effect of consonant position within the word on pharyngealization: word-initial and medial Cʰ had the same lingual configuration and the same extent of lingual displacement across all speakers.

3.4.2 Categorical vs. gradient assimilation (Question 2). In order to assess whether the regressive assimilatory effect of pharyngealization is gradient, the assimilated consonant (such as C1 in C1VCʰ2) must be compared to both plain C1 (C1VC2) and distinctive Cʰ (Cʰ1VC2) in terms of the tongue shape at the given regions and the magnitude of any lingual movement. If the assimilated consonant and the distinctive Cʰ are equivalent in terms of the tongue configuration, location of constriction (as indexed by the greatest displacement of a particular lingual region compared to other regions), and the degree of posterior constriction, this would indicate that the

process is categorical. Otherwise, if there is a discrepancy in these three criteria between the two consonants, the process would be considered gradient. I will first introduce the results bearing on the overall tongue shape in assimilated consonants before proceeding to the presentation of the results on the magnitude of the tongue root retraction.

The SS-ANOVA results which illustrated differences in tongue shape demonstrated that, at 50% of the phone or closure, the assimilated /t s/ before a pharyngealized consonant in CVC^s sequences were always articulated with the tongue root more posteriorly displaced than in plain /t s/ in plain context (CVC). The assimilated consonants had tongue configurations qualitatively similar to distinctive C^s (in C^sVC sequence), with a retracted tongue root, low and flat tongue body and low and retracted dorsum. This is an accordance with the hypothesis of a categorical assimilatory effect in terms of the tongue shape and location of the greatest lingual displacement.

The magnitude of tongue root displacement (be it forward or backward) is taken as an index for assessing the degree of posterior constriction. Analyzing the two assimilated consonants /t/ and /s/ together (at 50% of the phone or closure) demonstrated some differences between them in terms of the extent of assimilatory tongue root retraction; therefore, the results of the two consonants are presented separately.

3.4.2.1 Assimilation in /s/. The assimilated /s/ in an sVC^s sequence was articulated with the same lingual configuration and with a comparable extent of tongue root retraction and body lowering as in distinctive /s^s/ (Figure 3.14). A similar categorical effect was observed across the majority of speakers (60% of them) as indicated by an overlap between their SS-ANOVA tongue fits.

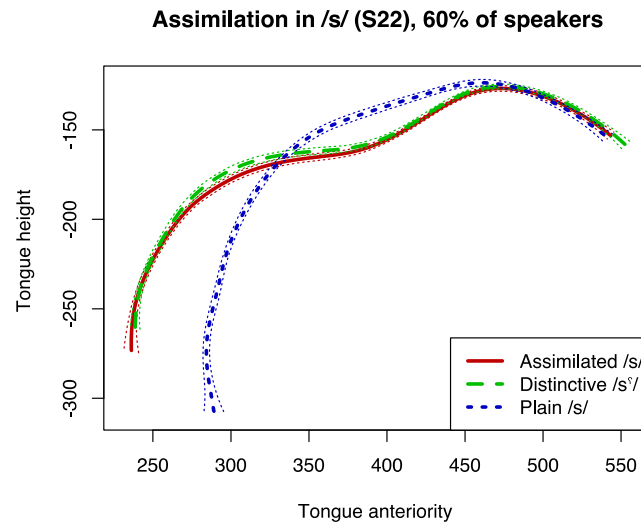


Figure 3.14. SS-ANOVA fit at mid-phone and 95% CI showing categorical assimilation in /s/.

These findings were consistent with the ratio and ANOVA results presented in the Appendix. For example, as illustrated in Figure 3.15, the categorical assimilation case presented above exhibited a large significant difference between assimilated /s/ and plain /s/ ($F(1,14) = 203.1$, $p < 0.001$, indicating the presence of an assimilatory effect) and a small insignificant difference from distinctive /sʰ/ ($F(1,38) = 0.46$, $p > 0.001$, reflecting categorical assimilation). Also, the retraction ratio of the normalized radius at the tongue root of assimilated /s/ was around 1, indicating full assimilation, compared to the retraction ratio of /s/ in a plain context which was around 0.

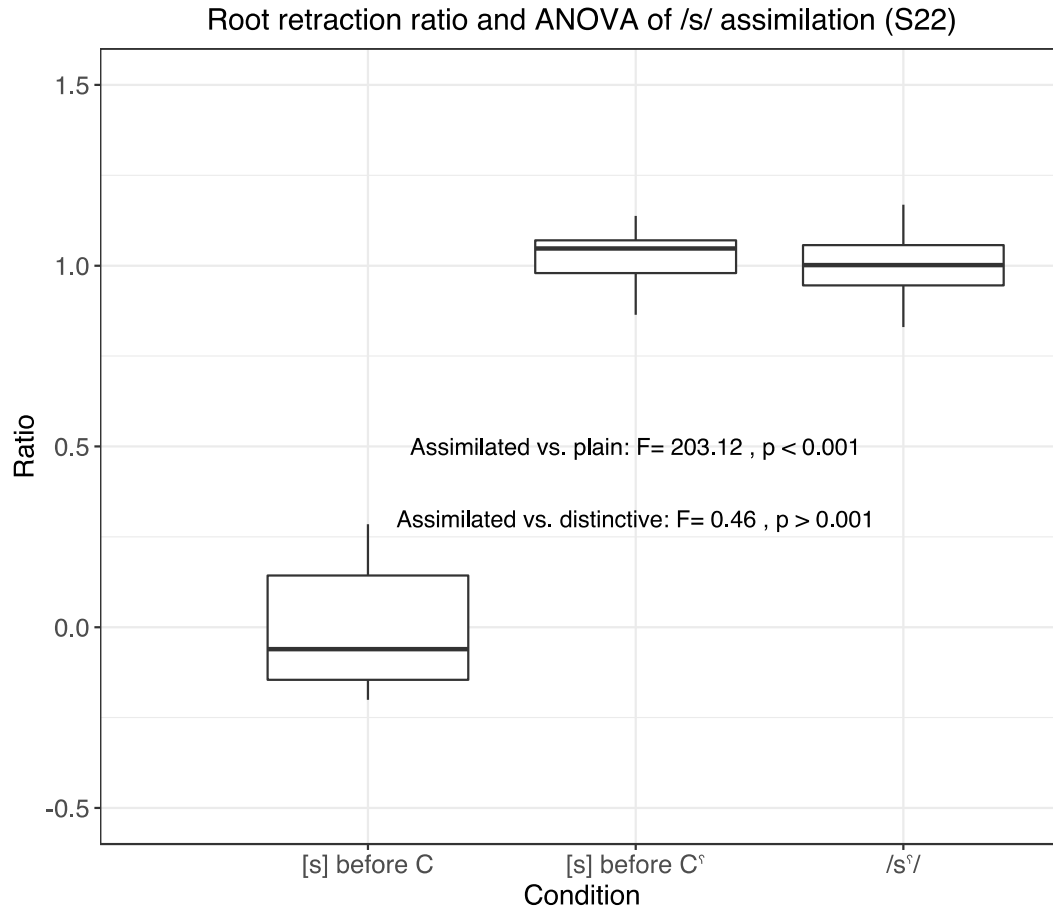


Figure 3.15. ANOVA and root retraction ratio of mean root radius showing categorical assimilation in /s/.

There was no consistent pattern across the remaining 40% of speakers. One speaker, S16, demonstrated a resistant /s/ based on the SS-ANOVA tongue spline of /s/ which was not significantly different from plain consonants (i.e. no assimilation) and on the ANOVA result of the difference between /s/ in plain and in pharyngealized context: $F(1,14) = 10.1$, $p > 0.001$. Another speaker, S02, showed a gradient effect on /s/ whose SS-ANOVA tongue fit was significantly different from both plain /s/ and distinctively pharyngealized /s^s/. This was consistent with the ANOVA results: the assimilated /s/ was significantly different at the tongue root from plain /s/ ($F(1,15) = 75.3$, $p < 0.001$) and distinctive /s^s/ ($F(1,34) = 39.3$, $p < 0.001$) with large F-

values. The other four speakers (who are not homogeneous in terms of the regional varieties of Eastern Peninsular Arabic they speak) exhibited no difference at the tongue root and other regions between the SS-ANOVA tongue fits of the assimilated and plain /s/ and distinctive /s^h/, which is surprising.

Based on acoustic evidence of vowel coarticulation in Jongman et al. (2011), one would expect assimilation to become smaller near the onset of the consonant (i.e. as the distance from the pharyngealization source increases). This, however, was not found in this study. The lingual configuration and the extent of the tongue root displacement was identical at all time points (frames) of the assimilated /s/ for all speakers (those who exhibited categorical assimilation and those who did not). The tongue splines of the assimilated /s/ at 25, 50 and 75% of its duration overlapped, indicating a consistent pattern throughout the consonant (Figure 3.16).

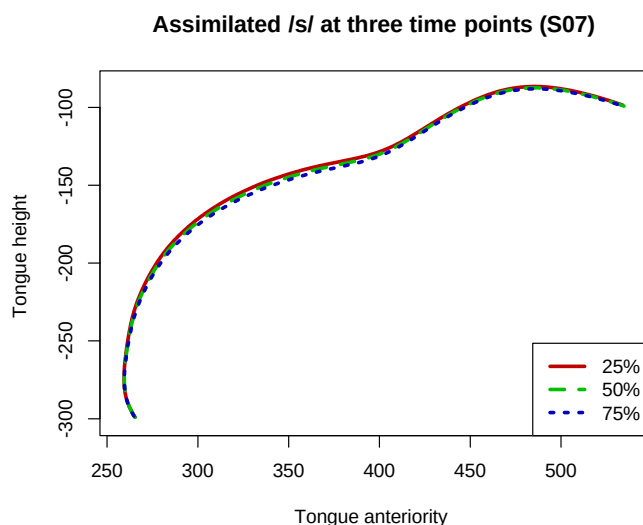


Figure 3.16. SS-ANOVA fit showing assimilated /s/ at three selected time points.

3.4.2.2 Assimilation in /t/. The other target consonant /t/ in tVC^s sequence also showed a categorical effect for the majority of speakers (80% of them) that is manifested as either complete resistance or categorical assimilation.

An example representative of 53.3% of speakers who exhibited /t/ resistance is shown in Figure 3.17 where the SS-ANOVA tongue fits of the consonant /t/ at the root area were not significantly different from plain consonants, indicating resistance to assimilation. For this resistance case, as demonstrated in Figure 3.18, the root retraction ratio of /t/ in pharyngealized context was very close to the ratio of /t/ in plain context. This is also indicated in the ANOVA results which showed that resistant /t/ was not different from plain /t/ at the tongue root ($F(1,34) = 1.8, p > 0.001$, reflecting no assimilatory effect) but significantly different from distinctive /tʰ/ ($F(1,55) = 396.3, p < 0.001$).

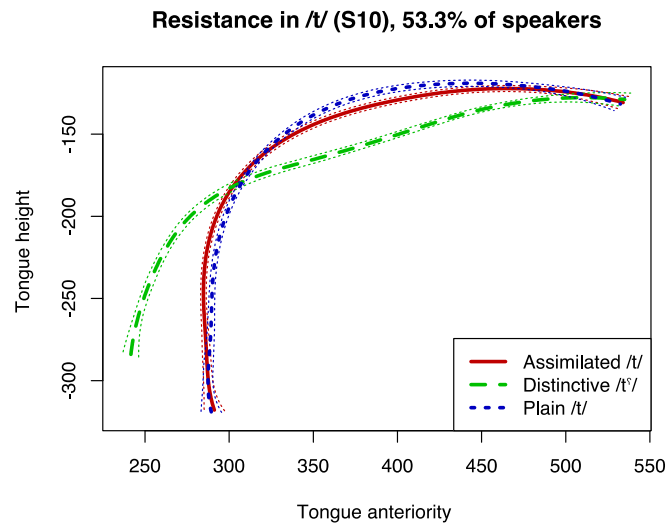


Figure 3.17. SS-ANOVA fit at mid-closure and 95% CI showing resistance in /t/.

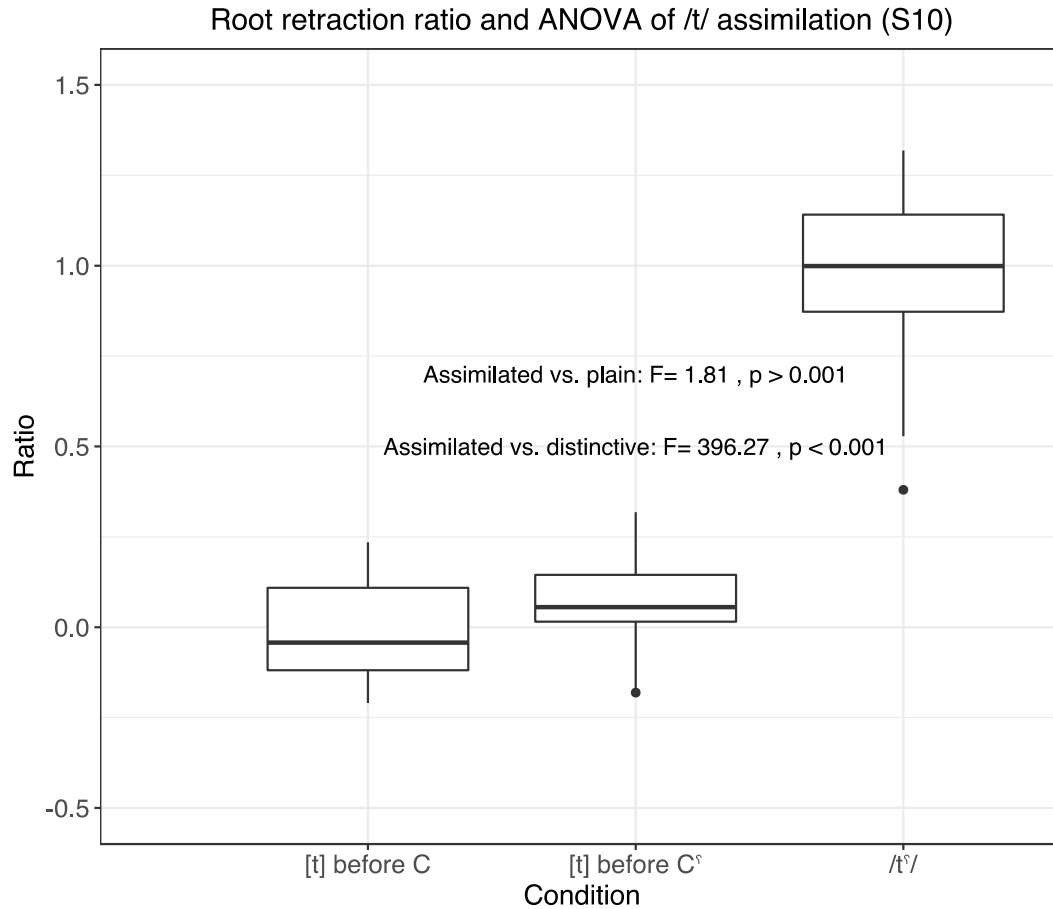


Figure 3.18. ANOVA and root retraction ratio of mean root radius showing resistance in /t/.

An example of full assimilation of /t/ representing 26.7% of speakers is shown in Figure 3.19 where the SS-ANOVA tongue fits of the consonant /t/ at the root area were not significantly different from distinctive C^s, indicating categorical assimilation. This categorization is consistent with the ANOVA results. For example, as shown in Figure 3.20, the categorical assimilation case presented here showed a large significant difference between assimilated /t/ and plain /t/ at the tongue root ($F(1,37) = 114.9, p < 0.001$, reflecting the presence of an assimilatory effect) and small insignificant difference between assimilated /t/ and distinctive /t^s/ ($F(1,55) = 8.8, p > 0.001$, reflecting a categorical assimilation). Consistent with these ANOVA results, the root retraction

ratio of assimilated /t/ was close to 1 and completely different from the root retraction ratio of /t/ in a plain context whose ratio was 0.

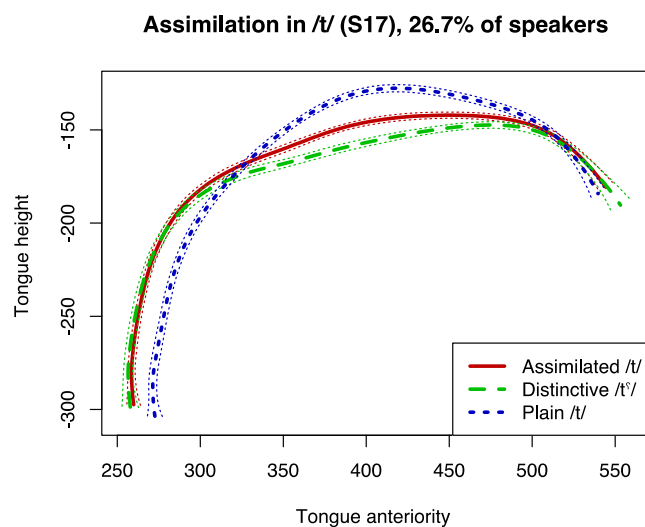


Figure 3.19. SS-ANOVA fit at mid-closure and 95% CI showing assimilation in /t/.

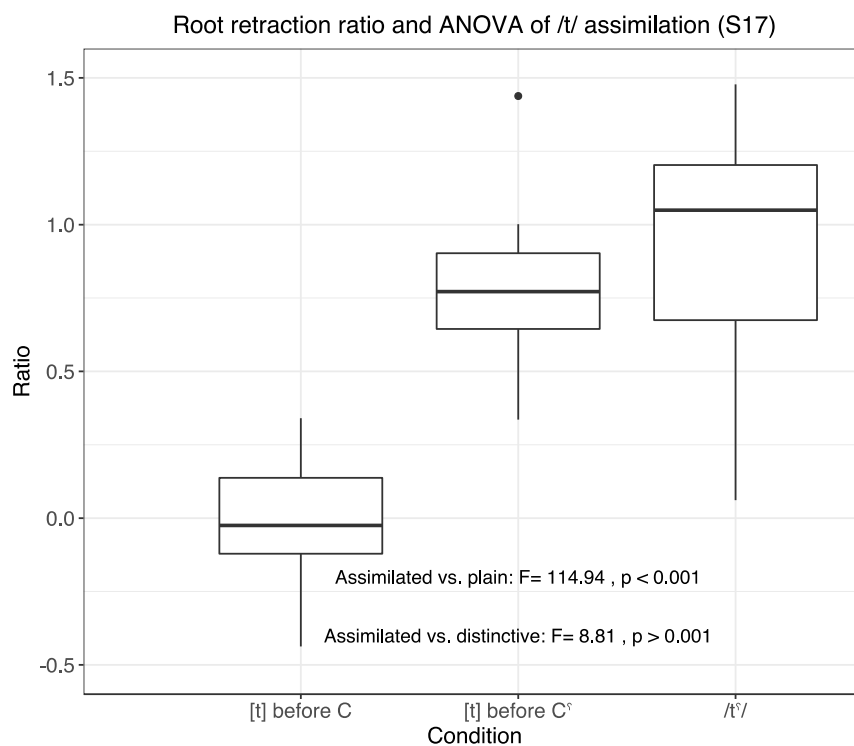


Figure 3.20. ANOVA and root retraction ratio of mean root radius showing categorical assimilation in /t/.

The assimilatory effect on /t/ for the remaining three speakers was different from the major trend. Although coarticulated /t/ was closer to either plain /t/ or to distinctive /tʰ/, their coarticulated /t/'s were classified as gradient according to categorization based on the $F=30$ threshold. For example, based on ANOVA, the coarticulated /t/ by speaker S18 was relatively different from plain /t/ ($F(1,33) = 43, p < 0.001$) but largely different from distinctive /tʰ/ ($F(1,53) = 314.5, p < 0.001$), suggesting that coarticulated /t/ was further close to plain /t/.

Despite the nature of the effect on /t/, the lingual configuration and the extent of the tongue root displacement was identical at the mid-closure and at the release of the assimilated /t/ for all speakers. The tongue splines of the assimilated /t/ at these time points overlapped, indicating a consistent pattern throughout the consonant (Figure 3.21).

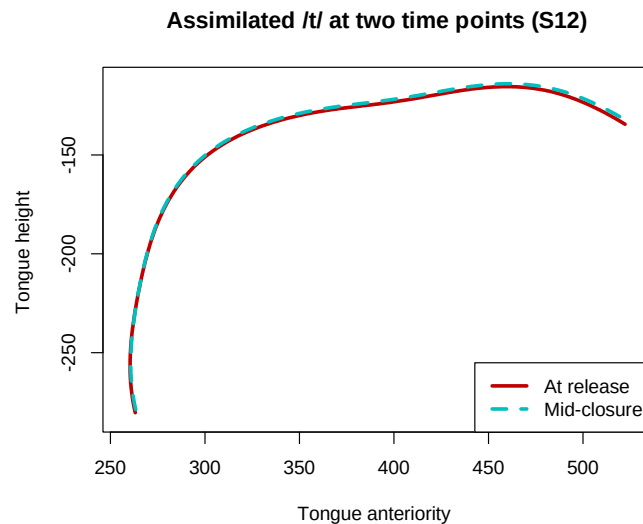


Figure 3.21. SS-ANOVA fit showing assimilated /t/ at two selected time points.

Two important conclusions can be drawn based on the examination of /t/ and /s/ in pharyngealized contexts. The assimilatory effect tended to be categorical across the majority of speakers and the categoricalness observed in both consonants was manifested as either a

substantial assimilation or a complete resistance. The fricative /s/ tended to be categorically assimilated most of the time (by 60% of speakers) and was never resistant, whereas the stop /t/ was resistant by many speakers (53.3%) and categorically assimilated by some (26.7%). It should also be noted that these differences are not speaker-based: a given speaker may exhibit resistance in /t/ but full assimilation in /s/.

3.4.2.3 Tongue root and body differences in /s t/ assimilation by speaker. A summary of the difference between SS-ANOVA fits of /s/ and /t/ at the tongue root and body portions for all speakers is presented in Figure 3.22 (the root retraction ratio and ANOVA of all speakers can be found in the Appendix). Unlike the scatterplots in §3.4.1 and as explained earlier (in §3.3.5.3.1), the y-axis here represents the difference between tongue fits at the tongue body area (75° - 105°). Comparing the SS-ANOVA fits of the assimilated /s/ to plain /s/ in a plain context indicated that the assimilated /s/ had a more retracted root than /s/ in a plain context by all speakers (positive root difference in Figure 3.22 right). The degree of root retraction and tongue body lowering in assimilated /s/ were not different from distinctive /s^ɣ/ by most speakers (Figure 3.22 left where x and y values are around 0). Speaker variation was more evident in the assimilated /t/ in which the root assumed a position close to that of distinctive /t^ɣ/ by some speakers and distant from it by other speakers (Figure 3.22 left). The diagonal angle of the ellipsis reflects a positive correlation between tongue root and body position: the greater the root retraction in C^s compared to assimilated C, the greater its tongue body lowering. This was evident for the two portions of the tongue in /s/ vs. /s^ɣ/ and /t/ vs. /t^ɣ/.

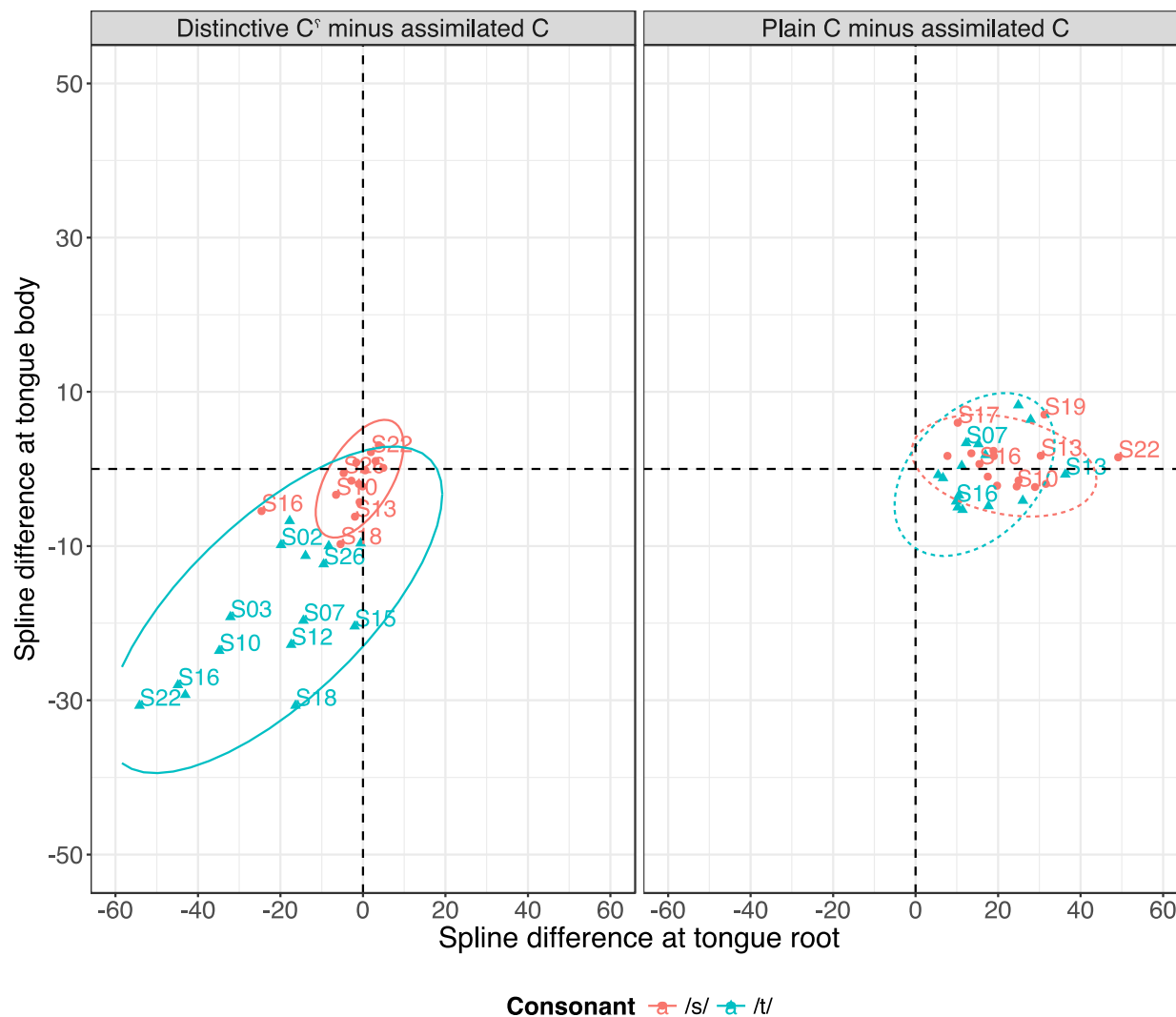


Figure 3.22. The difference between SS-ANOVA fits comparing assimilated /t s/ to distinctive /tʰ sʰ/ (left) and to plain /t s/ (right) for all speakers. Negative root difference in the left means the assimilated C is less retracted than distinctive Cʰ, and positive difference in the right means assimilated C is more retracted than plain C.

3.5 Discussion

The purpose of the present study was twofold: to examine the tongue configuration in phonemic emphasis and to assess the categoricalness of its contextual effect on the preceding consonant, two controversial issues among researchers.

3.5.1 Lingual configuration in phonemic pharyngealization. The present study supports the view of the speakers of Eastern Peninsular Arabic articulated emphatics as pharyngealized which is similar to that found in other dialects ([E. Al-Tamimi & Heselwood, 2011](#); [J. Al-Tamimi, 2017](#); [Ali & Daniloff, 1972](#); [Laufer & Baer, 1988](#); [Zeroual & Clements, 2015](#); [Zeroual et al., 2011](#)) rather than uvularized ([Al-Ani, 1970](#); [Al-Solami, 2017](#); [Bin-Muqbil, 2006](#); [Zawaydeh, 1997, 1998, 1999](#); [Zawaydeh & de Jong, 2011](#)). The conflicting results obtained in previous studies could be due to dialectal differences ([Khattab et al., 2006](#)), but the survey in Table 1.2 reveals that the disagreement among articulation studies cannot be attributed to dialectical or methodological differences. Rather, there is variation in the production of emphasis and it could be a speaker-level variation based on evidence that studies of the same dialect (e.g. Palestinian Arabic) found both pharyngealization ([Laufer & Baer, 1988](#)) and uvularization ([Herzallah, 1990](#)).

From the articulation of the Eastern Peninsular Arabic speakers, it seems clear that emphasis is a form of pharyngealization: the key articulator responsible for producing it is the tongue root, which is substantially retracted towards the posterior pharyngeal wall, thus constricting the pharyngeal tube horizontally. Due to this root movement and given the hydrostatic nature of the tongue, the tongue mass is pulled posteriorly and lowered. The tongue dorsum moves downward along with the tongue body which also takes on a flat shape for most speakers, as compared to its convex shape in non-pharyngealized consonants. This configuration differs from uvularization, in which the dorsum is pulled upward and backward. More specifically, the dorsum

depression and the tongue root retraction are the key differences between the uvularized consonants reported in some studies and the pharyngealized ones examined in other studies including this dissertation.

This configuration in producing pharyngealization is consistent across all speakers (with no systematic gender-based difference), in the two phonemically pharyngealized consonants /tʰ/ and /sʰ/, at the two positions (word-initially and medially), and at all time points during the consonant duration. Although not directly measurable with ultrasound, other articulatory correlates may result from the retraction of the tongue root. According to ([Esling, 1996, 2005](#)), there is a physiological contingency between tongue root retraction and larynx elevation which contributes to constricting the pharyngeal tube even further. The retraction of the epiglottis is another possible concomitant movement associated with the root retraction: although the epiglottis can retract independently from the root, moving the tongue root necessarily entails epiglottal movement. Although these articulatory movements were not directly examined in the present ultrasound experiment, they could contribute to the pharyngeal quality of the consonants under investigation.

3.5.2 Categoricalness of assimilatory pharyngealization. Another major finding of this study is that the regressive contextual effect of /tʰ sʰ/ is categorical, suggesting that the process is phonological rather than phonetic. The categoricalness is assessed according to three criteria: the tongue shape, the location of maximal lingual displacement and the degree of this displacement.

The three factors yield similar results which vary according to the identity of the assimilated consonant. Plain /s/, when preceding a pharyngealized consonant /tʰ/ or /sʰ/, is pharyngealized and produced by 60% of speakers with the same tongue contour and extreme tongue root retraction as in phonemically pharyngealized consonants. It is the tongue root that undergoes the maximal lingual displacement in both assimilated /s/ and phonemically

pharyngealized consonants, and the assimilated consonant is produced with the same extent of posteriority as the phonemic C^s . This categorical assimilation, in terms of the tongue shape as well as the magnitude of retraction, is constant at the three time points of the fricative – at 25, 50 and 75% of its duration. Note however, that, as discussed in the results section, an important proportion of speakers (40%) exhibit a variety of other patterns.

The other target consonant /t/ is either categorically assimilated or unaffected by pharyngealization (with very few exceptions of coarticulation). In cases of assimilation, the tongue root approaches a position similar to that of phonemically pharyngealized consonants and in the other cases, its position is similar to /t/ in a plain context. The number of speakers who exhibited resistance in /t/ exceeded the number of those who assimilated, reflecting a greater tendency of this consonant to resist assimilation.

CHAPTER 4

Phonological Factors and Assimilation

4.1 Introduction

Some assimilatory effects can result in an output with a great deviation from its canonical or underlying form leading, in turn, to a reduction in its articulatory, acoustic and perceptual distinctiveness. Other effects, however, are associated with less substantial deviations, thus maintaining the perceptual and articulatory distinctiveness of the phone.

It has been hypothesized by some researchers that phonological factors such as phonemic contrast can restrict the assimilatory effects that segments exert upon each other. According to [Manuel \(1987, 1990, 1999\)](#), ‘output constraints’ can predict how much tolerance a given segment has to deviation from its ideal distinctive characteristics as a result of coarticulation. Contrast has been found to be a predictor of vowels’ susceptibility to the assimilatory influence of neighboring vowels ([Manuel & Krakow, 1984](#)) and of adjacent nasal consonants ([Barlaz et al., 2018](#); [Delvaux et al., 2008](#); [Scarborough et al., 2015](#)). Other evidence, however, suggests that V-to-V coarticulation and nasalization ([Farnetani, 1990](#); [Solé, 1995](#)) are independent of contrast. Articulatory investigation of the interaction between contrast and other contextual variations such as assimilatory pharyngealization has not previously been carried out.

4.2 Research Question and Hypothesis

The purpose of this chapter is to study whether (and to what extent) phonemic contrast plays a role in constraining assimilatory pharyngealization in Eastern Peninsular Arabic. Out of the full set of plain coronal consonants in this language /t d n r θ ð s z l/, only four are contrastive with pharyngealized counterparts: /t d ð s/ vs. /t̤ d̤ ð̤ s̤/. The ultrasound study reported here aims to examine the articulatory correlates of two sets of coronals: those that have pharyngealized counterparts, such as /t s/, and those that do not, such as /n l/.

Evidence from the experiment reported in the previous chapter has already shown that the two coronals that contrast in pharyngealization had a categorical behavior: /s/ was categorically assimilated, whereas /t/ either fully resisted or underwent categorical assimilation in most cases. The fact that categorical assimilation is a possible outcome of pharyngealization in these consonants is already an indication that contrast preservation is not the driving force in itself but that the actual underlying principle may have to do with structure preservation.

The questions that remain to be addressed in order to have all the elements needed to solve this puzzle are whether (and why) the presence of phonemic contrast leads to smaller and less gradient coarticulation than its absence. Such questions are addressed by comparing the pharyngealization effect in the two sets of coronals: the coronals that contrast in pharyngealization (e.g. /t s/) versus those that are non-contrastive (e.g. /n l/).

If contrast is the driving force behind the categorical effect, consonants that do not bear a pharyngealization contrast should not strictly exhibit categorical effects (i.e. complete assimilation or resistance), but should rather exhibit gradient coarticulation. Given that categoricalness is assessed by the magnitude of the assimilatory pharyngeal gesture and its location, control consonants /n l/ should undergo a smaller tongue root displacement (and probably have a less

posterior constriction) than the consonants with a pharyngealization contrast, /t s/, when they exhibit categorical assimilation. Under this hypothesis, if the resistance found in /t/ is due to contrast, the other coronals may not exhibit the same resistance. If, on the other hand, /t s/ and /n l/ undergo the same categorical effect (i.e. the same magnitude of tongue root retraction or resistance), we would have evidence that contrast is not responsible for the categorical behaviour of /t s/.

4.3 Methodology

The wordlist used to address this question consisted of 28 words organized in (near)-minimal pairs contrasting in pharyngealization in which the source consonant was either phonemically pharyngealized /t^s s^s/ or plain /t s/. The target consonant that preceded the source consonant was a plain coronal with either a pharyngealized counterpart (i.e. /t s/) or not (i.e. /n l/). The first set of coronals /t s/ will be called ‘contrastive’ and the control ones ‘non-contrastive’ to denote their phonemic status with respect to pharyngealization. The inter-consonantal vowel intervening between the source and target consonant was always [a a:]. The complete wordlist containing these two contrast conditions is given in Table 4.1. The word selection criteria were the same as those used in the previous chapter: prosodic, phonetic and phonological factors were controlled.

Data for this part of the study were collected in the same ultrasound recording session as the data presented in the previous chapter. Mid-sagittal ultrasound images (along with audio recordings) were generated from 15 adults using Ultraspeech 1.2. Images were recorded while participants read the target words embedded in the same frame sentence as in Chapter 3. Each word was repeated five times by each participant. The data were subjected to the same processing procedures as in Chapter 3.

Two frames were selected from the stops /t t^ɕ/ (at mid-closure and at release) and three frames from non-stops /s s^ɕ n l/ (at 25, 50 and 75% of consonant duration). This resulted in a total of 9,570 ultrasound images: 3,652 frames for words containing contrastive /t s/ and 5,918 for words containing non-contrastive /n l/.

Table 4.1

Complete list of words used to address the effect of phoneme contrast

Target pharyn. contrast	Target	Source	Example	Gloss
Contrastive	/t/	/t/	/tatli:f/	‘ruining’
		/t ^ɕ /	/tat ^ɕ fi:f/	‘cheating in weight’
		/t/	/tatbi:l/	‘using spices’
		/t ^ɕ /	/tat ^ɕ bi:l/	‘drumming’
		/s/	/tasafi:r/	‘travelling-PL’
		/s ^ɕ /	/tas ^ɕ afi:r/	‘whistling-PL’
		/s/	/tasmi:m/	‘poisoning’
		/s ^ɕ /	/tas ^ɕ mi:m/	‘design’
	/s/	/t/	/satra:t/	‘covering-FEM’
		/t ^ɕ /	/sat ^ɕ ra:t/	‘hitting on face-FEM’
		/t/	/satarha:/	‘covered her’
		/t ^ɕ /	/sat ^ɕ arha:/	‘hit her on face’
Non-contrastive	/n/	/t/	/zanat/	‘committed adultery-FEM’
		/t ^ɕ /	/zanat ^ɕ /	‘strangled-MS
		/t/	/na:teʒha:/	‘her/its result’
		/t ^ɕ /	/na:t ^ɕ egha:/	‘pronouncing her’

Target pharyn. contrast	Target	Source	Example	Gloss
		/s/	/na:sfa:t/	‘destroying-FEM-PL’
		/s ^ɕ /	/na:s ^ɕ fa:t/	‘fair-FEM-PL’
		/s/	/naserha:/	‘her eagle’
		/s ^ɕ /	/nas ^ɕ erha:/	‘her victory’
	/l/	/t/	/salat/	‘entertained-FEM’
		/t ^ɕ /	/salat ^ɕ /	‘salad’
		/t/	/xalat/	‘passed-FEM’
		/t ^ɕ /	/xalat ^ɕ /	‘mixed-MSc’
		/s/	/əxtəlas/	‘robbed-MSc’
		/s ^ɕ /	/əxtəlas ^ɕ /	‘redeemed-MSc’
		/s/	/tala:senha:/	‘her quarrelling’
		/s ^ɕ /	/tala:s ^ɕ egha:/	‘her adhesion’

4.4 Results

Since the two control coronals /n l/ and test coronals /t s/ were not equally sensitive to pharyngealization across speakers, SS-ANOVA results are presented for each consonant separately. After assessing whether they undergo assimilation and the extent of the effect, the four consonants will be compared to each other, both in plain and in pharyngealized context, to examine whether there is any difference between /t s/ and /n l/ in their assimilatory pattern that could be attributable to contrast. Although more focus will be placed on the tongue shape at the mid-point of the consonant or closure (i.e. at 50%) of a given speaker, results at other time points will also be presented.

4.4.1 Assimilatory pharyngealization in individual consonants.

4.4.1.1 Contrastive coronals /t s/. The two coronals /t s/, which are contrastive for pharyngealization, exhibited a categorical effect induced by pharyngealized consonants, as shown in the previous chapter.

The fricative /s/ was produced with the same extent of tongue root retraction as in phonemically pharyngealized consonants by the majority of speakers (60% or 9/15). As in Figure 4.1 (left), /s/ in a pharyngealized context was identical to distinctive /s^h/ (shown in red) but different from plain /s/ in plain context (shown in green).

The stop /t/ was categorically affected by the pharyngealization trigger: the tongue root assumes a position that is either similar to plain /t/ in 53.3% (8/15) of speakers or similar to distinctive /t^h/ in 26.7% (4/15) of speakers (based on the F=30 threshold).

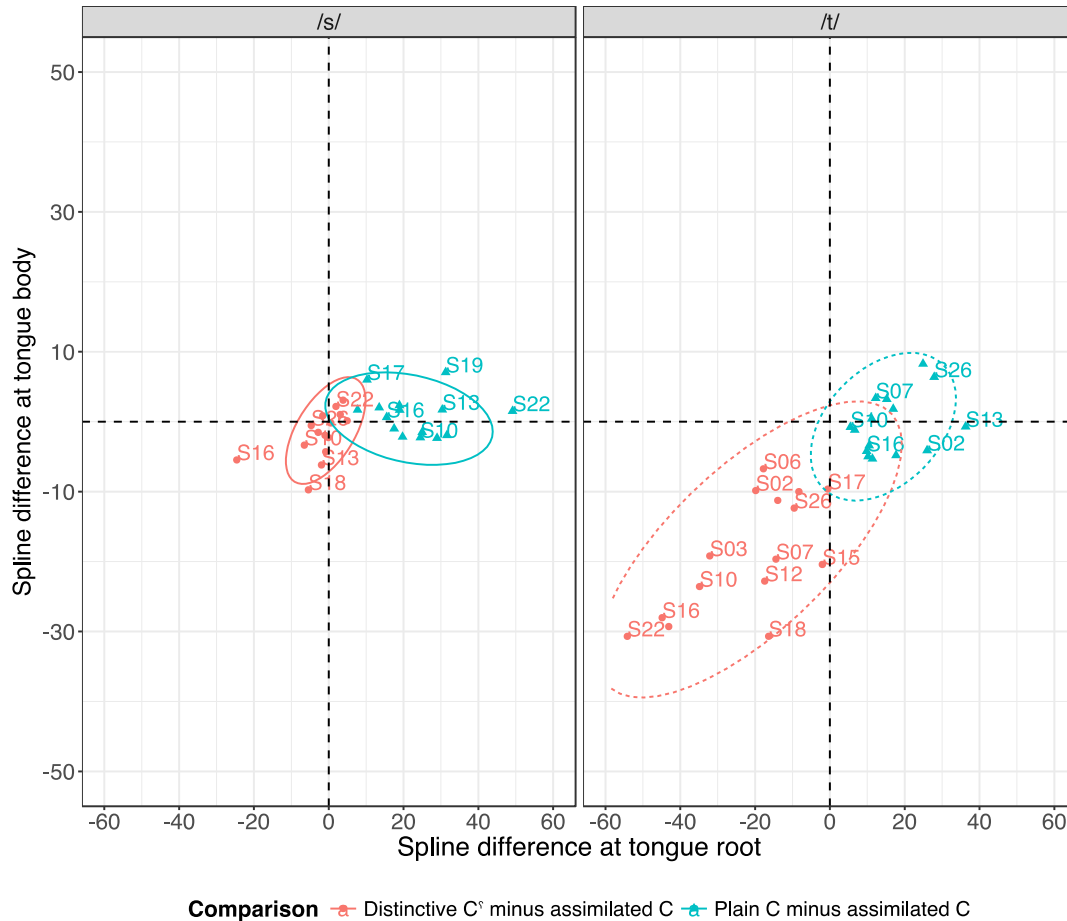


Figure 4.1. The difference between SS-ANOVA fits comparing assimilated /t s/ to distinctive /tʰ sʰ/ and to plain /t s/ for all speakers.

4.4.1.2 Non-contrastive coronals /n l/. Similar to /t s/, the consonants /n l/, that are not contrastive in pharyngealization, demonstrated different assimilatory patterns.

4.4.1.2.1 coarticulation in /n/. When occurring in a pharyngealized context, the nasal coronal /n/ was resistant to coarticulation in most speakers (93.3% or 14/15) with two of them exhibiting marginal resistance. At the midpoint of the consonant, /n/ (as in /nasʰerha:/ ‘her victory’) was produced with a tongue root position close to that of /n/ in a plain context (as in /naserha:/ ‘her eagle’), suggesting a lack of coarticulatory effect. This is presented in Figure 4.2.

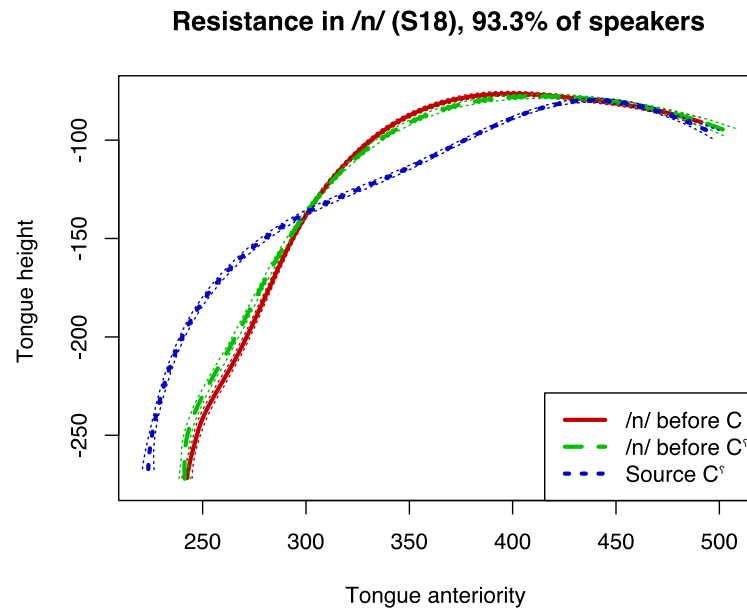


Figure 4.2. SS-ANOVA fit at mid-phone and 95% CI showing resistance in /n/.

This categorization of /n/ as resistant was based on the tongue root retraction ratio and ANOVA of the root radius. The root retraction ratio of /n/ in a pharyngealized context for the speaker shown above is very close to the ratio of /n/ in a plain context (the median of both was around 0 - 0.1 as in Figure 4.3), suggesting that the normalized radius at the tongue root area was similar in both contexts and that the source C^s had no effect on the root radius of /n/. This was also reflected in the ANOVA of /n/ in the two contexts ($F(1,37) = 3.43, p > 0.001$) and in the ANOVA of /n/ in a pharyngealized context versus source C^s ($F(1,87) = 162.4, p < 0.001$), with the later ANOVA indicating that the root position in resistant /n/ was not retracted as in source C^s.

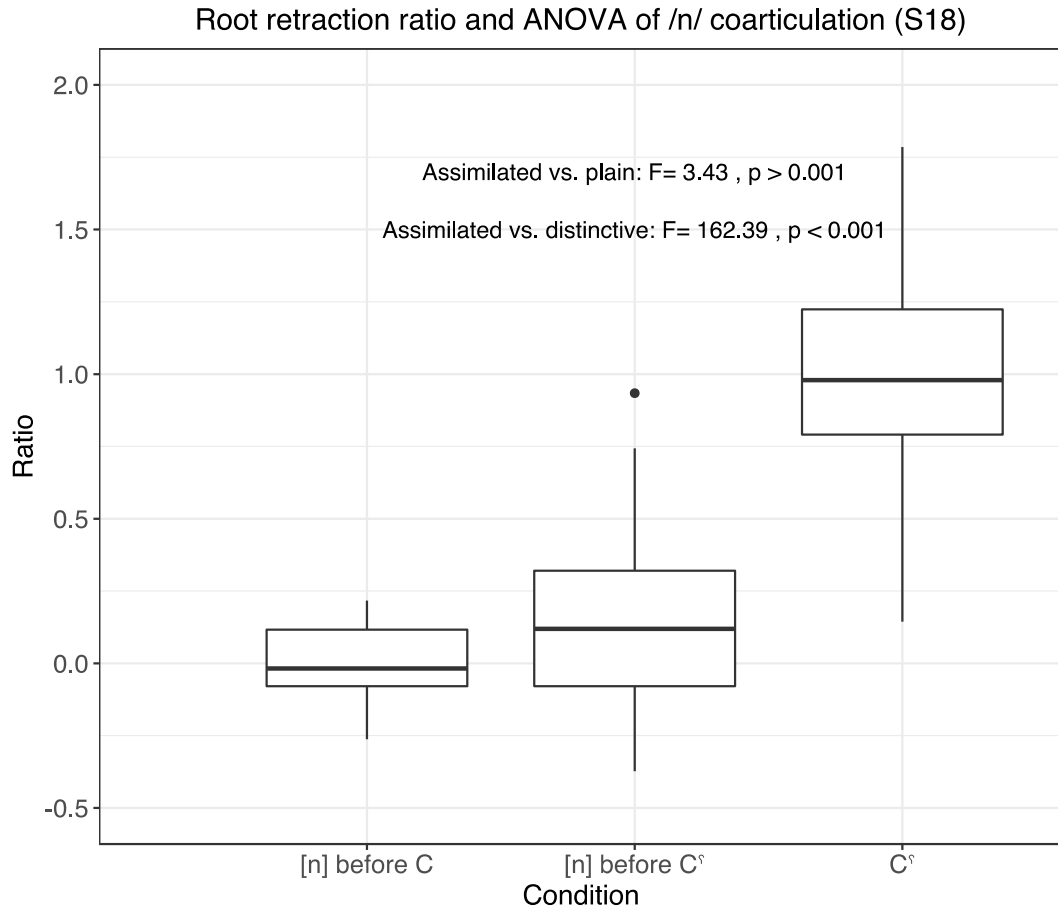


Figure 4.3. ANOVA and root retraction ratio of mean root radius showing resistance in /n/.

This pattern applied to 93.3% of speakers; however, the ANOVA results of two of them (S16 and S15) showed that the difference between /n/ in the two contexts was larger than the rest of speakers (S16: $F(1,32) = 29.6$, $p < 0.001$; and S15: $F(1,21) = 29.5$, $p < 0.001$). Their SS-ANOVA showed that the tongue root of coarticulated /n/ was slightly retracted compared to /n/ in a plain context, but further forward from its position in source C^s. The nasal by these two speakers is nonetheless classified as (marginally) resistant according to the ANOVA threshold.

SS-ANOVA of the remaining speaker, S17, showed an overlap at the tongue root between the three tongue fits: the fits of /n/ in the two contexts and the fit of source C^s. This also converged with the ANOVA of /n/ in both contexts ($F(1,35) = 4.08$, $p > 0.001$) and ANOVA of coarticulated

/n/ versus source C^s ($F(1,85) = 24.1$, $p < 0.001$), both of which indicating a small difference between coarticulated /n/ and the two reference points.

The tongue shape of all the speakers were similar at the three time points of /n/ that were examined, which indicates consistent patterns throughout the entire consonant in a pharyngealized context (Figure 4.4). The speaker whose SS-ANOVA is presented in Figure 4.4 was one of those who produced a resistant /n/ (as seen in the position of the tongue root $x \approx 300$). The time difference between the frames corresponding to these three time points was only 16-17ms in most instances of /n/ (i.e. consecutive frames), which may explain the overlap of tongue splines.

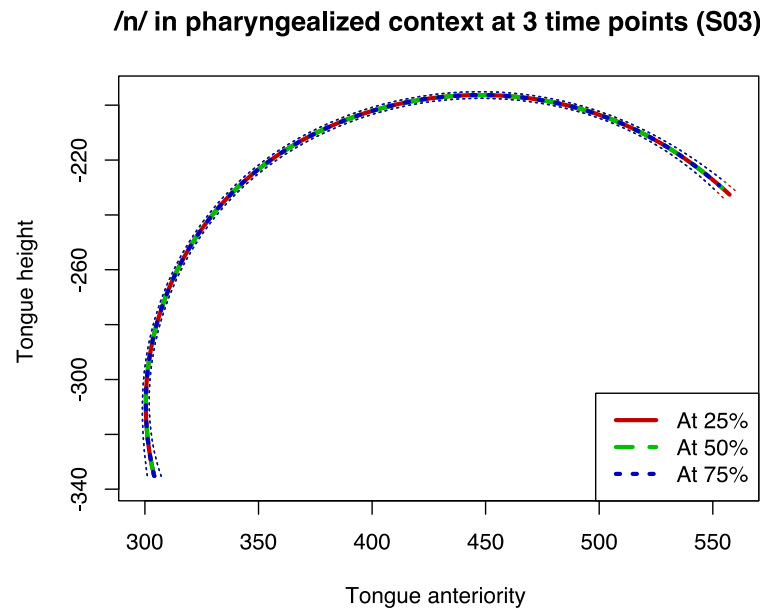


Figure 4.4. SS-ANOVA fit of /n/ in pharyngealized context at three selected time points.

4.4.1.2.2 *Assimilation in /l/*. Unlike /n/, the lateral coronal /l/, which is also not contrastive in pharyngealization, was always affected by the neighboring pharyngealized consonant (e.g. /xalat^s/ ‘mixed’) compared to /l/ in a plain context (e.g. /xalat/ ‘passed’). However, the magnitude

of the assimilatory effect varied considerably among speakers. It was categorical for some speakers (46.7% or 7/15) and gradient for others (40% or 6/15 speakers).

For those who exhibited a categorical effect, the assimilatory retraction was very extreme to the extent that the root position was not different from the source C^s as presented in Figure 4.5. The tongue fits of assimilated /l/ and source C^s completely overlapped at the root area for some of these speakers, and this root retraction in assimilated /l/ was accompanied with a lowering at the tongue body area.

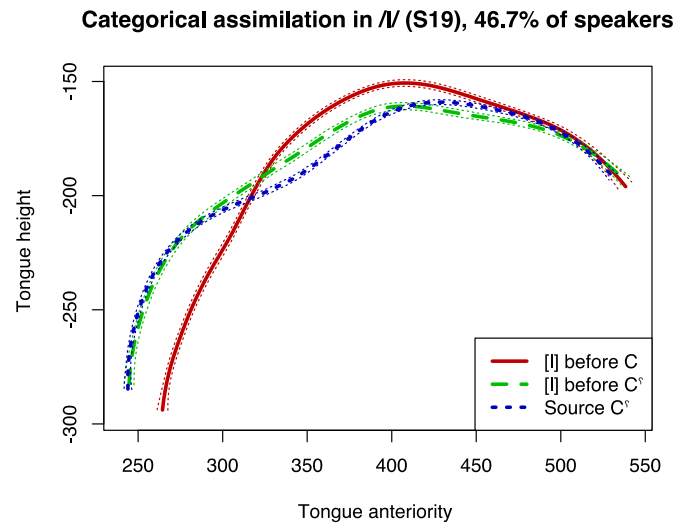


Figure 4.5. SS-ANOVA fit at mid-phone and 95% CI showing categorical assimilation in /l/.

The root retraction ratio of assimilated /l/ in these cases was high and close to 1, as evident in the boxplot in Figure 4.6 which presents the retraction ratios for the same speaker whose SS-ANOVA is shown above. The difference between the root radius of /l/ in pharyngealized and in plain context for this speaker was large and significant ($F(1,35) = 486.23$, $p < 0.001$), which indicates that the following pharyngealized consonant causes a significant retraction in /l/. The fact

that the root radius of assimilated /l/ was not different from that of source C^s ($F(1,82) = 1.45, p > 0.001$) suggests that the lateral was categorically retracted.

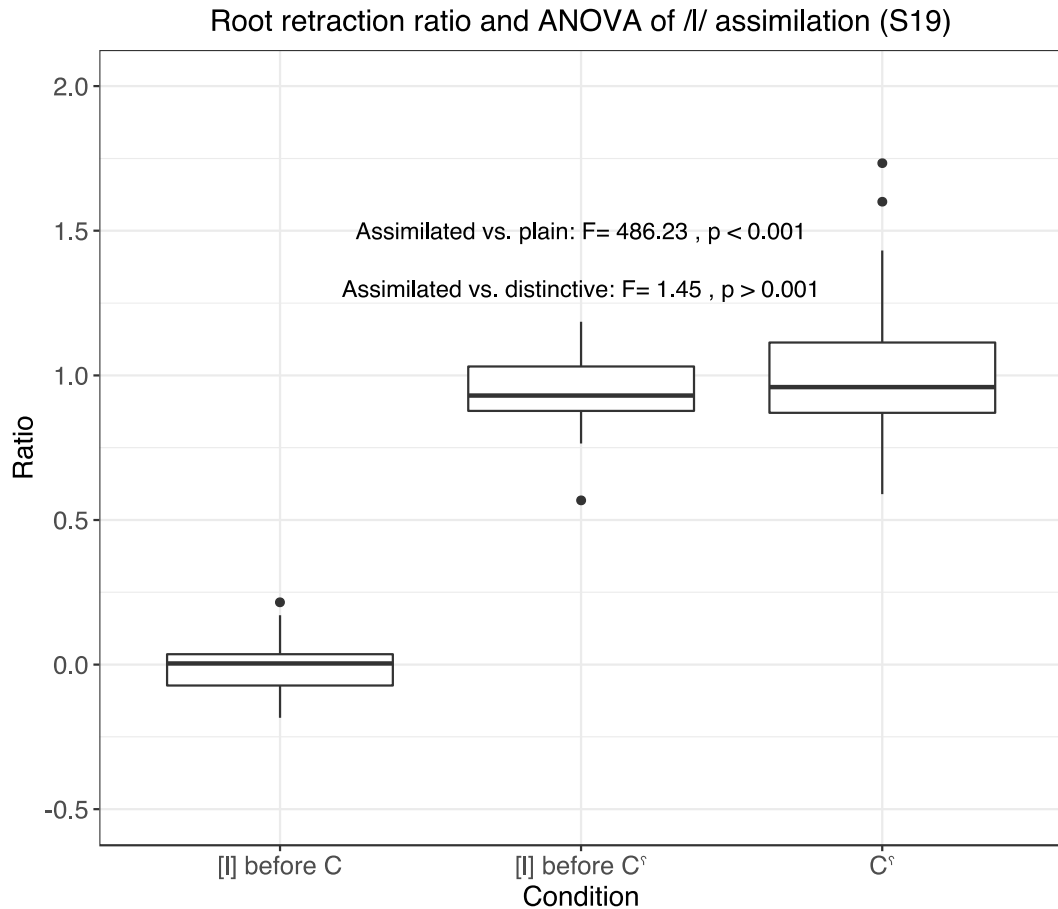


Figure 4.6. ANOVA and root retraction ratio of mean root radius showing categorical assimilation in /l/.

The magnitude of coarticulatory tongue root retraction in /l/ was less extreme in another 40% (6/15) of speakers, and the effect was therefore categorized as gradient in these cases. An example speaker of this category is presented in Figure 4.7 where the tongue fit of coarticulated /l/ is more retracted at the root area than the fit of /l/ in a plain context, but less retracted than the fit of source C^s. The tongue body is also lower than in plain /l/ but not as low as in source C^s.

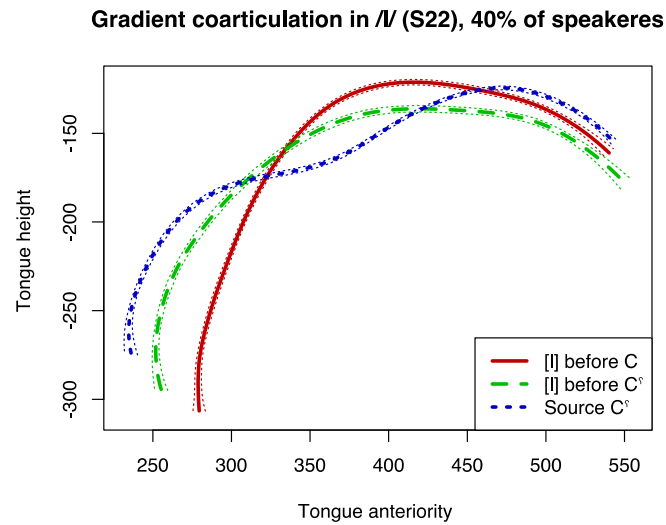


Figure 4.7. SS-ANOVA fit at mid-phone and 95% CI showing gradient coarticulation in /l/.

The categorization of the coarticulated /l/ of those 40% speakers as gradient was based the ratio and ANOVA of the radius at the tongue root. The root retraction ratio of coarticulated /l/ for the speaker shown in Figure 4.7 and 4.8 was between the ratio of /l/ in a plain context and source C^s : the median of the coarticulated /l/ ratio was around 0.6 compared to a ratio of 0 for /l/ in a plain context and of 1 for the source C^s . ANOVAs on the root radius of coarticulated /l/ and plain /l/ yielded an F-value comparable to that of coarticulated /l/ versus source C^s , suggesting that the root in coarticulated /l/ was between the two extreme positions. The difference in root radius with plain /l/ was $F(1,38) = 103.12$, $p < 0.001$ (suggesting a significant retraction in coarticulated /l/), while it was $F(1,88) = 121.48$, $p < 0.001$ with source C^s , suggesting that the coarticulatory retraction in /l/ was less extreme than the phonemic retraction in source C^s .

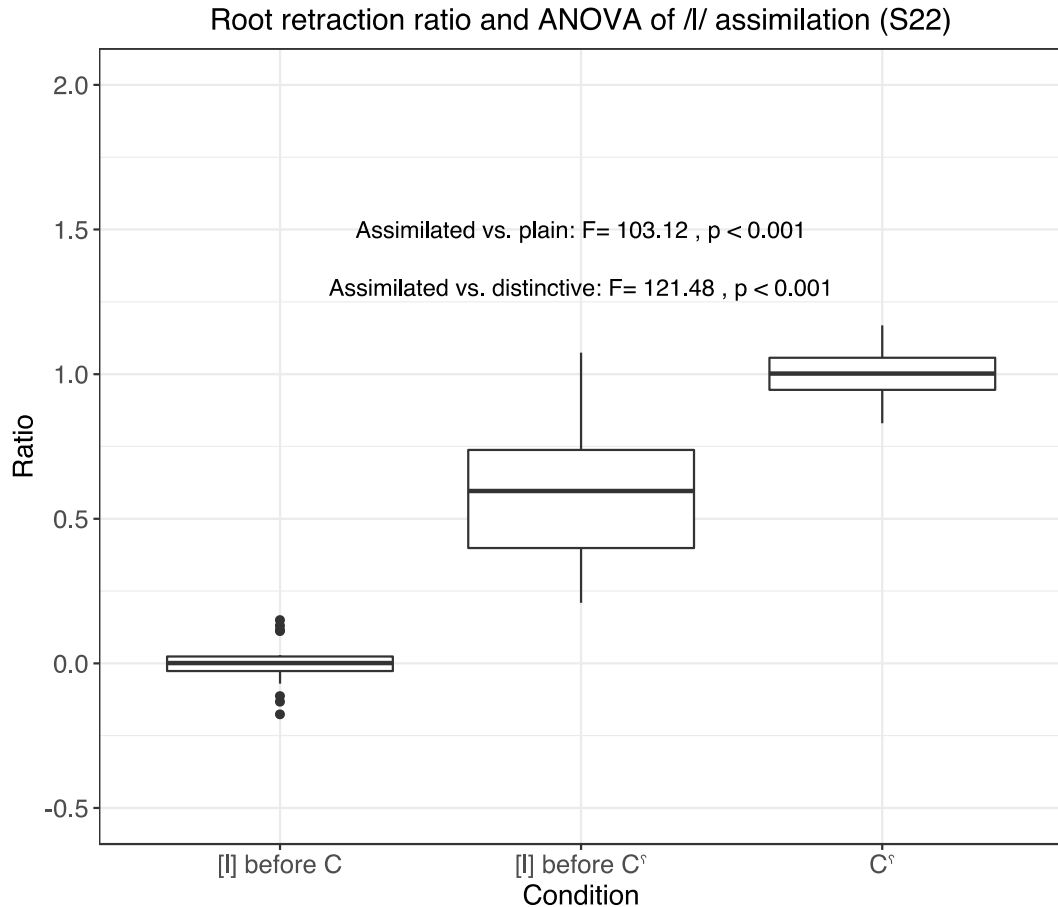


Figure 4.8. ANOVA and root retraction ratio of mean root radius showing gradient coarticulation in /l/.

The tongue configuration of /l/ in a pharyngealized context, whether retracted or not, was not only observed at 50%, but also at the 25% and 75% of the phone duration, indicating that the extent of assimilatory root retraction was constant throughout its duration (Figure 4.9). Worth noting here is that, the splines at the three time points overlapped throughout the entire tongue for some speakers; however, for other speakers such as the one presented in Figure 4.9, the overlap was observed at the tongue body, dorsum and root but not at the tongue blade. In fact, there was a slight increase in the blade height towards the end of the phone, which could be due to the fact that in some words, the lateral was preceded by the low vowel /a/. This means that there was a gradual increase in blade height from its low position in /a/ to a higher position required for producing /l/.

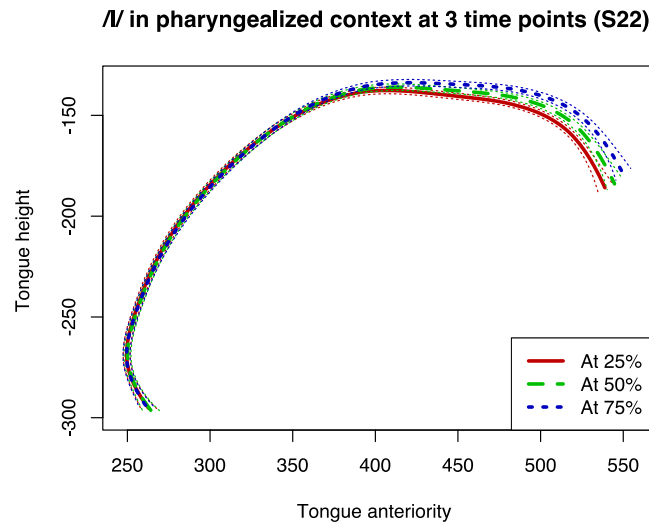


Figure 4.9. SS-ANOVA fit of /l/ in pharyngealized context at three selected time points.

4.4.1.2.3 Tongue root and body difference in /n l/ assimilation by speaker. To assess the extent of coarticulatory tongue root retraction in /n/ and /l/ for every speaker, the difference between the coarticulated /n l/ versus the same consonants in a plain context and the source C^s was calculated at both the tongue root ($0^\circ - 45^\circ$) and tongue body angles ($75^\circ - 105^\circ$) of the respective SS-ANOVA fits. The results are presented in the scatterplot in Figure 4.10.

As shown in the right scatterplot, the magnitude of root retraction in the coarticulated /l/ varied among speakers from a minimal retraction to a relatively large one, with the difference from plain /l/ reaching 40mm by some speakers. The extent of assimilatory retraction and the variation among speakers in /l/ was greater than in /n/. This speaker variation was consistent with the ANOVA classification showing two assimilatory patterns in /l/ (resistance and assimilation) but one in /n/ (resistance). Regardless of the degree of root displacement, the pharyngealized source induced little or no lowering in the tongue body in both /n/ and /l/ compared to their plain counterparts.

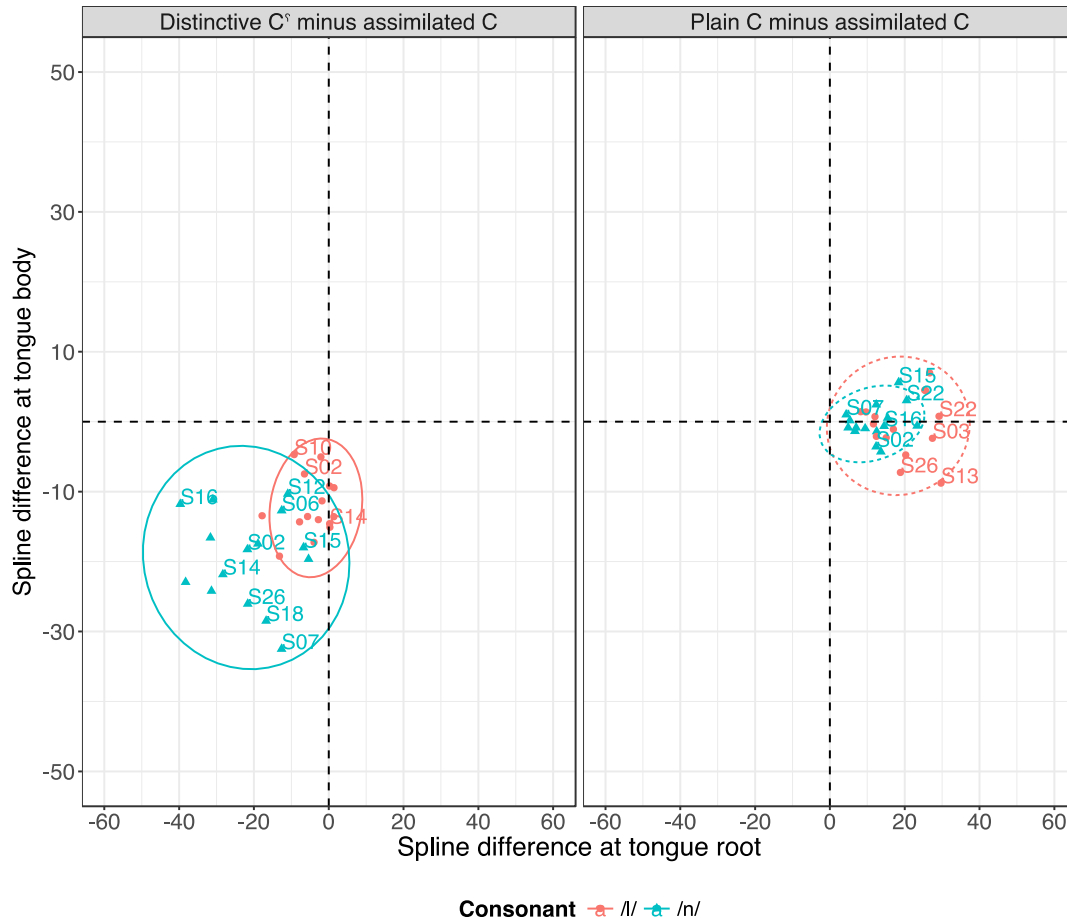


Figure 4.10. The difference between SS-ANOVA fits comparing assimilated /n l/ to distinctive C^s (left) and to plain /n l/ (right) for all speakers. Negative root difference in the left means the assimilated C is less retracted than distinctive C^s, and positive difference in the right means assimilated C is more retracted than plain C.

The left scatterplot in Figure 4.10 illustrates the magnitude of the root retraction of /l/ in relation to the pharyngealized source C^s to determine whether it is categorical or not.⁵ The difference in the root position of assimilated /l/ and the pharyngealized source C^s was around the

⁵ Note that the categorization of consonants (by speaker) as resistant, gradient or categorically assimilated was based on ANOVA results. The scatterplots help understand the distribution among speakers in terms of the size of the difference between the assimilated C and the two reference points (C in plain context and source C^s).

vertical zero line for some speakers, consistent with the ANOVA-based categorization of /l/ as categorical for these particular speakers. For other speakers, the root position in /l/ was less backward than in source C^s as indicated by the negative root difference in the left panel. Those speakers whose data points were further away from the vertical zero line in the left panel are categorized, according to ANOVA findings, as having gradient coarticulation in /l/. It is evident from the left scatterplot that the difference between the assimilated nasal /n/ and the source C^s was great for the majority of speakers. This means that speakers articulated /n/’s preceding a pharyngealized consonant with the tongue root displaced further forward in the pharyngeal cavity (i.e. less retracted) than its position in producing phonemic pharyngealization (as all data points were < 0). This is consistent with the categorization of /n/ as resistant to coarticulation according to ANOVA results.

4.4.2 Comparison of assimilation between contrastive and non-contrastive coronals.

Comparison of the four consonants in this experiment revealed that the consonants within the same contrast category (i.e. contrastive /t s/ and non-contrastive /n l/) demonstrated different assimilatory patterns. A summary of the observed assimilatory patterns per consonant are presented in Table 4.2. The [-continuant] consonants /t n/ exhibited a similar resistance pattern by at least some speakers: /t/ was resistant by 8/15 speakers and /n/ by 14/15 speakers. Likewise, [+continuant] ones /s l/ were categorically assimilated by some speakers regardless of whether they are contrastive in pharyngealization or not: /s/ was assimilated by 9/15 speakers and /l/ by 7/15 speakers.

Table 4.2

Assimilatory effects (by consonant) in pharyngealized context showing the number of speakers

	/t/	/s/	/n/	/l/
Resistance	8	1	14	0
Gradient coarticulation	3	1	0	6
Categorical assimilation	4	9	0	7
All consonants overlapping	0	4	1	2

4.4.2.1 Coronals in plain context. In a plain context, the four consonants varied in the exact position of the tongue root; however, this variation was not necessarily significant for all speakers (Figure 4.11). To illustrate this, the tongue root was positioned the furthest back in the pharyngeal cavity during the production of /l/ (e.g. /əxtəlas/ ‘robbed’) and the furthest front during /t/ (e.g. /tasmi:m/ ‘poisoning’) by 93.3% of speakers. The root was also retracted during the production of /s/ (e.g. /satarha:/ ‘covered her’) but to a lesser extent than in /l/ by 73.3% of speakers. While producing /n/ (e.g. /naserha:/ ‘her eagle’), the root overlapped with that of /t/ for half of speakers and was more posterior than that of /t/ for the other half (yet not significantly).

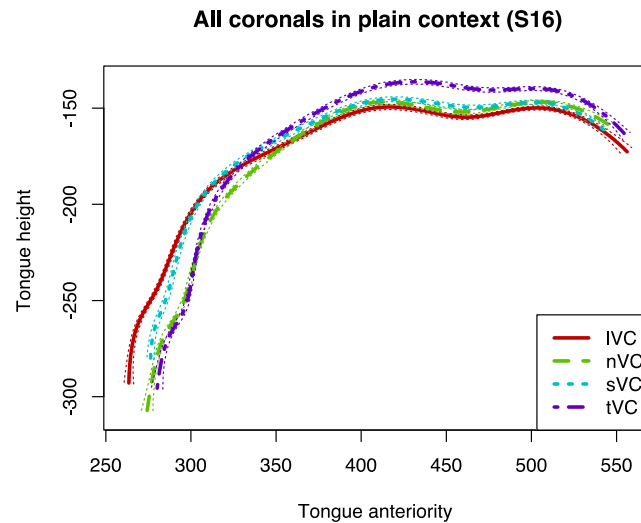


Figure 4.11. SS-ANOVA fit at mid-phone/closure and 95% CI showing (non)-contrastive coronals in plain context.

In other words, two obvious trends related to the root position emerged, as shown in the two scatterplots in Figure 4.12. First, the root position in non-contrastive /n/ and contrastive /t/ was not different, and the difference did not exceed 10mm for the majority of speakers as in Figure 4.12 (right). Second, the tongue root and body position in non-contrastive /l/ was similar to contrastive /s/ for most speakers, with /s/ being slightly less retracted than /l/ for some speakers (as seen in Figure 4.10 left).

The two pairs of consonants differed in tongue body height to a greater extent than in tongue root position: the majority of speakers produced /t/ with a higher tongue body than /n/ (negative difference at the tongue body). Also, /s/ was produced with a higher tongue body than /l/ by some speakers (negative y-value) and lower body by other speakers (positive y-value).

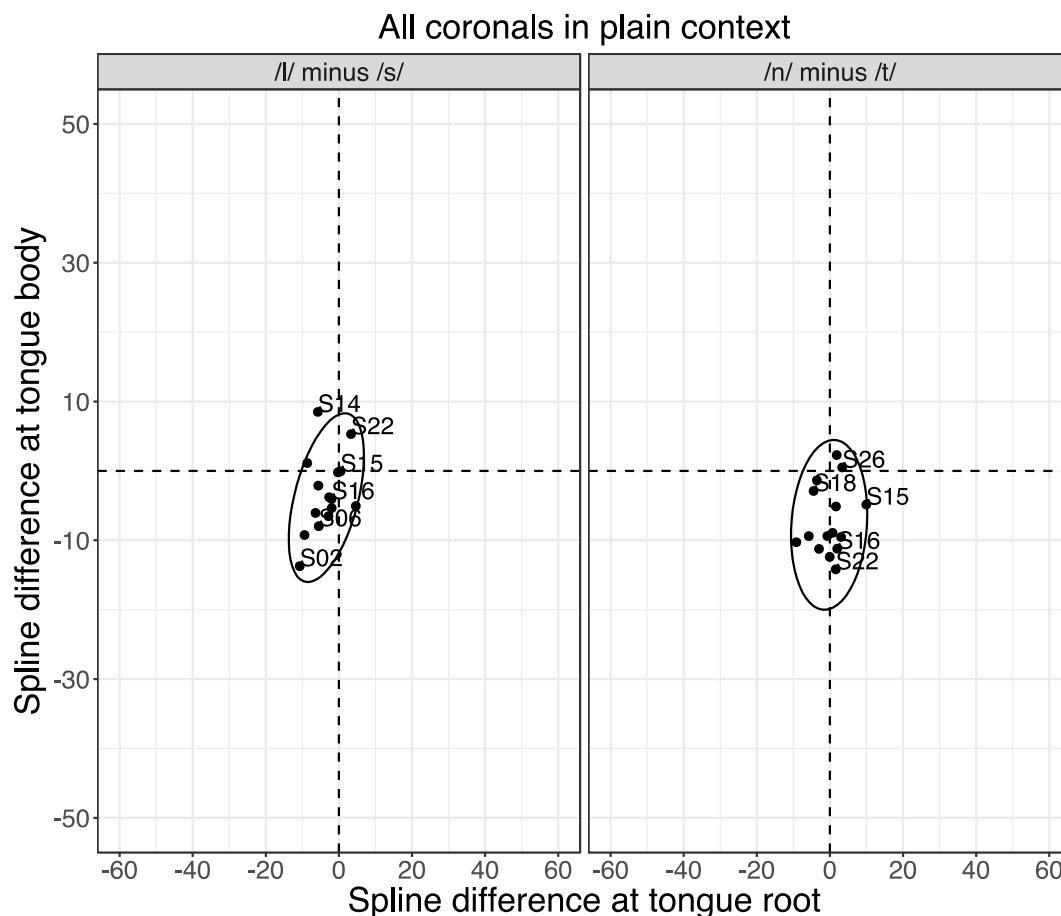


Figure 4.12. The difference between SS-ANOVA fits comparing each two coronals in plain context for every speaker. Negative root difference in the left means /s/ was less retracted than /l/ and in the right means /t/ was less retracted than /n/.

4.4.2.2 Coronals in pharyngealized context. Similar trends were also observed in the pharyngealized context as illustrated in Figure 4.13 (note that the speaker presented in this figure produced a gradient /l/, assimilated /s/ and resistant /n t/ according the ANOVA-based categorization). The largest retraction of the tongue root was during the production of /l/ (e.g. /əxtələs/ ‘redeemed’) for 86.7% of speakers, and the same extent of retraction was found in the production of /s/ (e.g. /satʰarha:/ ‘hit her on the face’) for 60% of speakers. The contrastive /t/ (e.g.

/tas^hmi:m/ ‘design’) had the least backward displacement of the root for 73.3% of speakers, a position that was not significantly different from that of /n/ (e.g. /nas^herha:/ ‘her victory’).

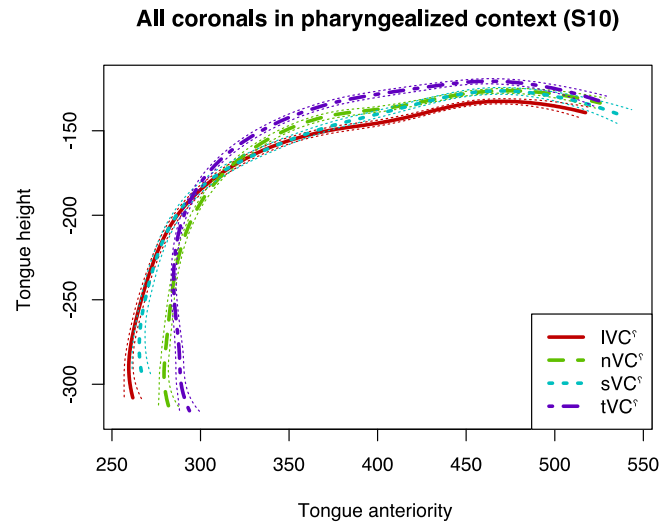


Figure 4.13. SS-ANOVA fit at mid-phone/closure and 95% CI showing (non)-contrastive coronals in pharyngealized context.

The scatterplots in Figure 4.14 summarize the difference between the tongue fits of the stops /n t/ and the consonants /l s/ for every speaker. The two consonants presented in the right scatterplot shared the resistance effect by some or most speakers: non-contrastive /n/ was resistant by 14/15 speakers and contrastive /t/ was resistant by 8/15 speakers. What this suggests is that the resistance in /t/ cannot be attributed to the pharyngealization contrast because the non-contrastive /n/ was also resistant to pharyngealization. However, the difference at the root between these two consonants varied in terms of its direction and magnitude. That is, it indicates that /t/ was produced with a more retracted root than /n/ by some speakers (positive root difference) or with a less retracted root by other speakers (negative root difference). The magnitude of the difference is due to the fact that /n/ was resistant by all speakers (except one) and that /t/ was resistant by some speakers and

categorically assimilated by others. Given this, the speakers who produced a resistant /n/ and assimilated /t/ had a large difference between the two consonants in terms of the root position.

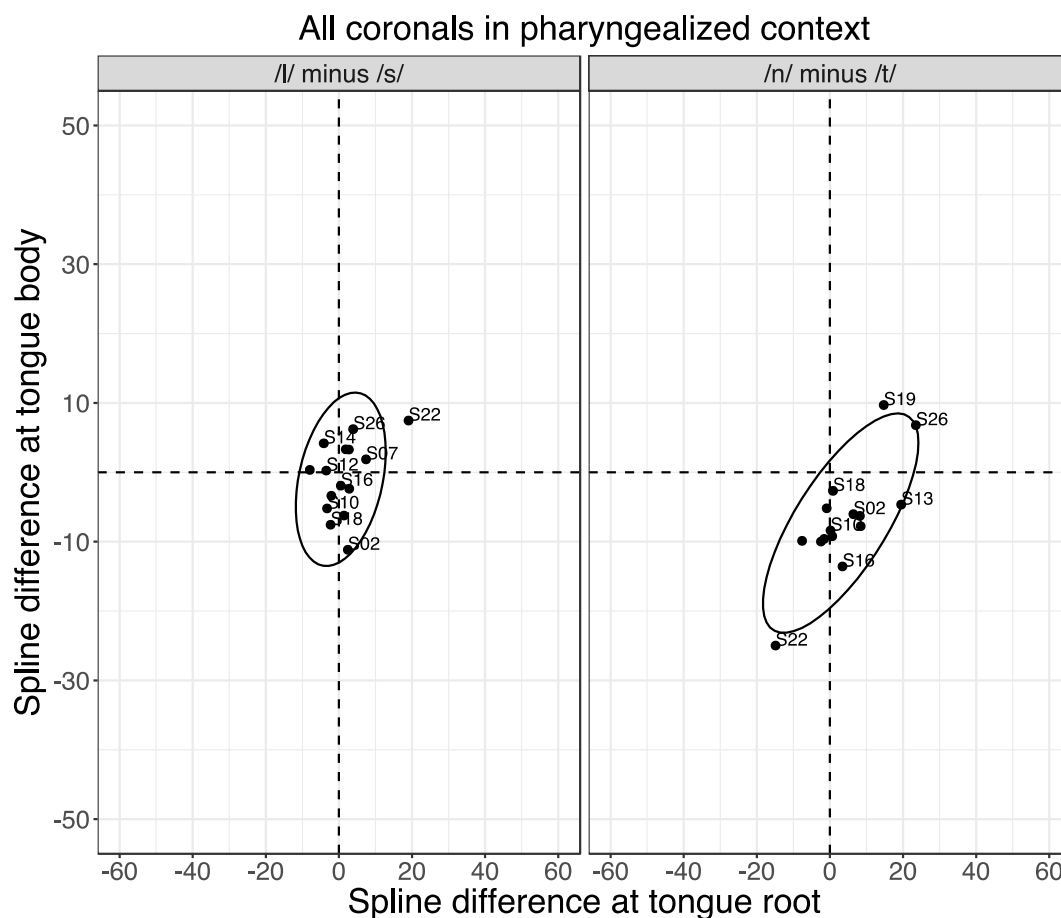


Figure 4.14. The difference between SS-ANOVA fits comparing each two coronals in pharyngealized context for every speaker. Negative root difference in the left means /s/ was less retracted than /l/ and in the right means /t/ was less retracted than /n/.

As presented in the left scatterplot, non-contrastive /l/ and contrastive /s/ occurring in a pharyngealized context were produced with a similar degree of root retraction and tongue body height, with the differences being less than 10mm for all speakers except one, as seen in Figure 4.14 (left). The equal distribution of speakers around the vertical and horizontal zero lines means that about half the speakers produced /l/ with more root retraction than /s/ and the other half

produced /s/ with more root retraction. Likewise, the number of speakers who produced /s/ with higher tongue body than /l/ was comparable to those who produced /l/ with a higher tongue body than /s/. The finding that root and body difference for a few speakers was at the zero lines indicates that the magnitude of root retraction and body lowering in /l/ and /s/ was the same in these cases. What these results suggest is that, although one of these two consonants is contrastive in pharyngealization and the other is not, both underwent comparable degrees of assimilatory root retraction.

4.4.2.3 Magnitude of assimilation in all coronals. The coarticulation pattern was assessed by measuring the difference between the tongue fit of the consonant in a plain context and the same consonant in a pharyngealized context. The difference was measured at the tongue root and body and was done for each speaker individually as usual. The results are presented in Figure 4.15. Using the root retraction ratio to compare the two consonant conditions provided a precise measure of the magnitude of assimilatory root retraction per consonant (for every speaker). I subtracted the median ratio of C in a plain context from the median ratio of C in a pharyngealized context. This subtraction, which was done for every consonant and speaker, supported two major findings presented in Figure 4.16 and 4.17, as explained below.

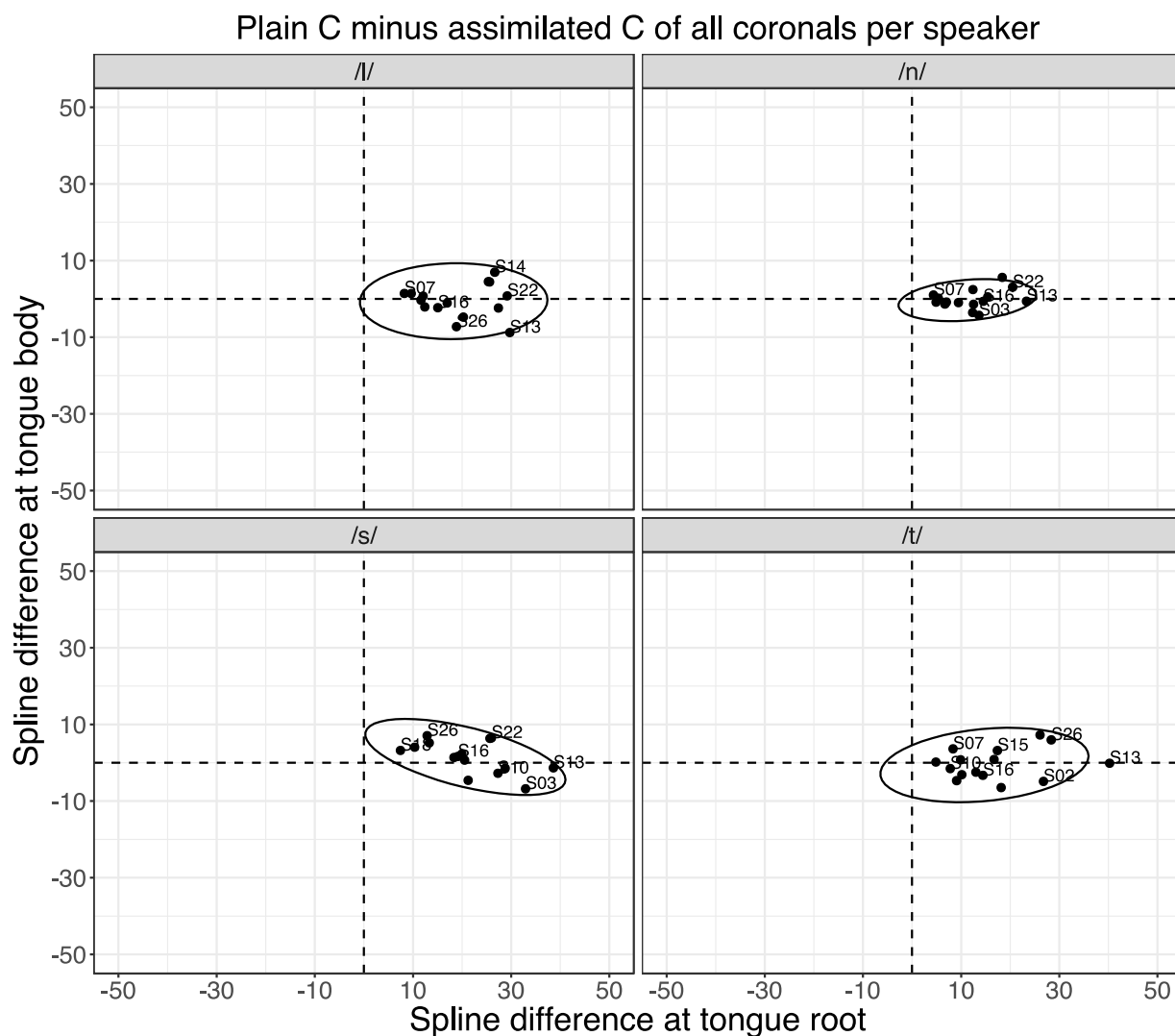


Figure 4.15. The difference between SS-ANOVA fits of all coronals comparing plain vs. assimilated C for every speaker. Positive root difference means that assimilated C was more retracted than plain C

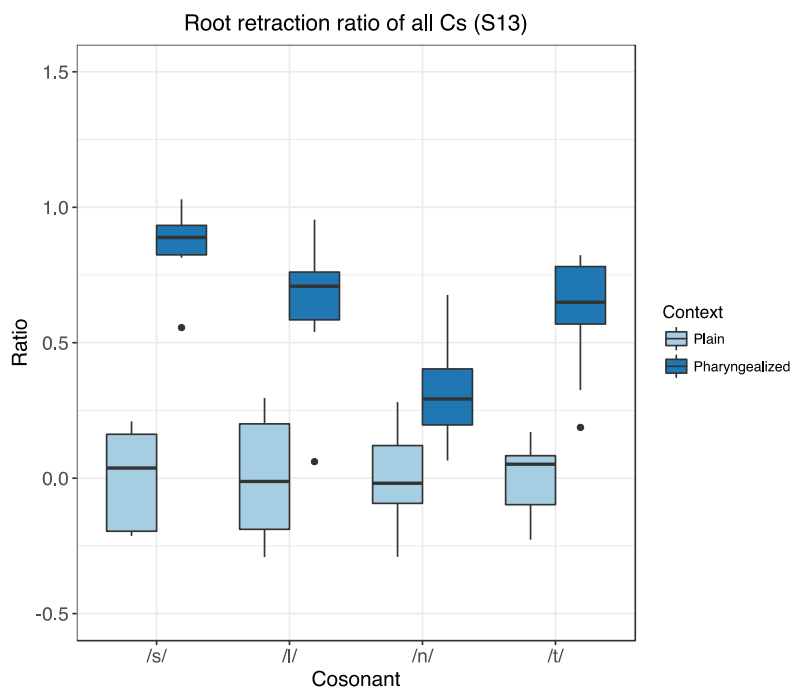


Figure 4.16. Root retraction ratio for the degree of assimilatory root retraction in /s l n t/, showing more retraction in /s/ than /l/ and in /t/ than /n/.

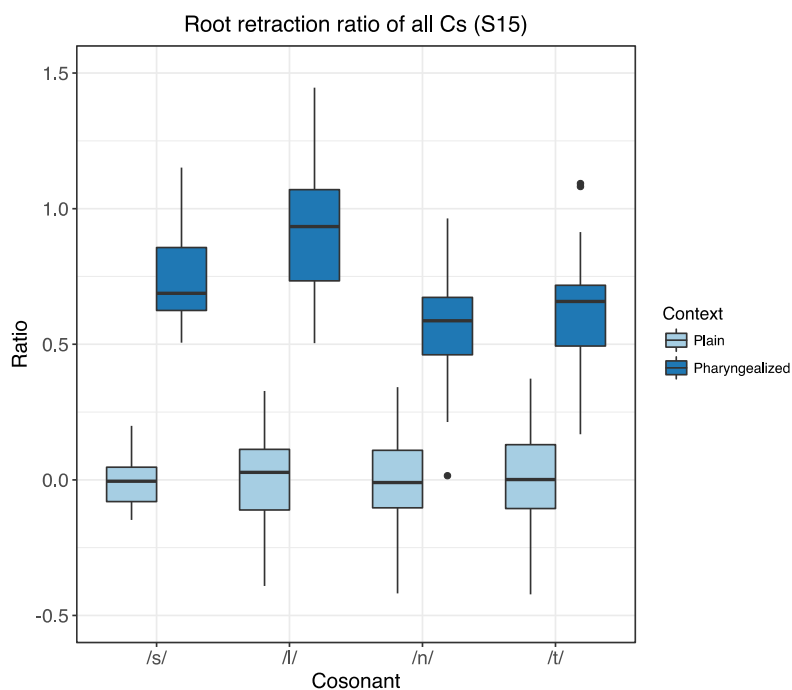


Figure 4.17. Root retraction ratio for the degree of assimilatory root retraction in /s l n t/, showing less retraction in /s/ than /l/ (although both were categorically assimilated).

The comparison of all four coronals in plain and pharyngealized context using SS-ANOVA fits and the retraction ratios (Figure 4.15, 4.16 and 4.17) confirms the findings in §4.4.2.2:

- The four consonants varied in terms of the size of the root difference between the plain and assimilation conditions as seen in the ratio plots (Figure 4.16, 4.17), and in the distribution of speakers around the vertical line in the scatterplot (Figure 4.15).
- Consistent with the finding that /n/ was resistant by 14/15 speakers and /t/ by 8/15 speakers, the scatterplot illustrates that the two consonants /t n/ were produced with a marginal root retraction (<10mm) by some speakers. Also, as shown in Figure 4.16, the majority of speakers produced /t/ with a greater root retraction than /n/, which is not surprising given the fact that not all speakers produced a resistant /t/ unlike /n/.
- Non-contrastive /n/ had the least speaker variation (Figure 4.15) and the minimum difference between the two conditions across speakers (Figure 4.16, 4.17).
- The ratio comparison shows that the assimilated /s/ incurred a greater assimilatory root retraction than assimilated /l/ by nearly half the speakers (8/15 speakers, Figure 4.16), and less retraction than /l/ by the other speakers (7/15 speakers, Figure 4.17). Of course, the size of the difference between the two consonants varied among speakers.
- Finally, there was no clear distinction between the four consonants in terms of assimilatory tongue body displacement (Figure 4.15). Speakers were distributed almost equally around the horizontal line (i.e. the line where tongue body difference = 0) in all four consonants.

4.5 Discussion

The investigation carried out in this experiment was motivated by the lack of consensus among the existing acoustic studies on the interplay between phoneme contrast and phonetic

variability. Leading to the postulation of Manuel's 'output constraints' hypothesis, some studies ([Manuel, 1987, 1990, 1999](#); [Manuel & Krakow, 1984](#)) argued that languages with a dense vowel inventory tend to allow less V-to-V coarticulation than languages with a sparse inventory. Although segments are subject to contextual variability, speakers should maintain the distinctiveness of a phoneme by restricting the magnitude of this variability. Another line of research reported no such effect of phonemic contrast on coarticulation ([Beddor et al., 2002](#); [Solé, 1995](#)).

Taking this disagreement into consideration and building on the findings of Chapter 3, the current study aimed to investigate whether contrast contributes to the categoricalness of assimilatory pharyngealization in /t s/ which are contrastive in pharyngealization. To address this issue, these consonants were compared to other coronals, e.g. /n l/, that do not have such a contrast. It was hypothesized that if contrast is responsible for the categoricalness of pharyngealization effects, the assimilatory patterns in non-contrastive coronals /n l/ should be less categorical than /t s/.

The findings reveal that, although they differ in pharyngealization contrast, all four except /n/ are susceptible to assimilatory pharyngealization but to varying degrees and with a great speaker variation. This variation means that a given consonant was assimilated or coarticulated by some speakers but is resistant by others. The three affected coronals /t s l/, differ in the spatial extent of the assimilatory displacement of the tongue root but not in the posteriority of the assimilated gesture. The overall trends are as follows:

- Contrastive /t/ is either categorically assimilated or resistant (and gradiently coarticulated by very few speakers);
- Contrastive /s/ is categorically assimilated;

- Non-contrastive /n/ is resistant; and
- Non-contrastive /l/ is either gradiently coarticulated or categorically assimilated.

Three general conclusions can be drawn based on these trends. First, phonemic contrast does not necessarily prevent coarticulation from taking place. This is based on evidence from the previous and the present chapters that consonants with pharyngealization contrast (e.g. plain /t s/) often undergo an assimilatory root retraction, which is the key gesture in producing pharyngealization. The fricative /s/ was assimilated by 60% of speakers and /t/ by 26.7% of speakers. Moreover, the assimilatory gesture in /t s/ was very extreme to the extent that, in many cases, the assimilated consonants were indistinct from their phonemically pharyngealized counterparts – e.g. the assimilated /s/ was produced with the same lingual configuration and the same (or similar) extent of root retraction as phonemically pharyngealized /s^ʕ/ – leading to a neutralization of contrast in these cases. Moreover, although /t/ is resistant to assimilation in some speakers, this resistance is not exclusive to consonants with pharyngealization contrast. Evidence comes from the finding that, similar to contrastive /t/, the non-contrastive /n/ is equivalently resistant to assimilatory pharyngealization by a subset of speakers, which rules out the hypothesis that contrast stimulates resistance in order to maintain phonetic distinctiveness.

Another important conclusion is that coronals with pharyngealization contrast tend to react categorically to pharyngealization induced by the neighboring consonants. The categorical outcome manifests itself as either resistance or full assimilation. With the exception of a few instances of /t/, the two contrastive coronals tend to avoid gradient coarticulation when occurring in a pharyngealized context. In fact, gradient coarticulation is more likely in consonants with no pharyngealization contrast: non-contrastive /l/ undergoes gradient coarticulation in 40% of

speakers whereas /t/ and /s/ tend to be categorically affected by pharyngealization (resistance and assimilation) in more than half the speakers (in /t/ by 80% of speakers and in /s/ by 66.7% of speakers). This partially supports the initial hypothesis that contrast enhances the categoricalness of assimilatory patterns among consonants with pharyngealization contrast.

Related to this is the finding that /t/ is more prone to resistance than /s/ and likewise, /s/ is more prone to full assimilation than /t/. The tendency of /s/ to fully assimilate more than /t/ is not surprising given the finding that, when both occur in a plain context, /s/ is produced with a more backward tongue root than /t/ by 73.3% of speakers (§4.4.2.1) which suggests that it is more susceptible to assimilatory root retraction than /t/. The same principle may also apply to the two non-contrastive coronals, /l/ and /n/. The former, when occurring in a plain context, tends to be articulated with a more backward position of the tongue root than /n/ in the same context (as explained in details in §4.4.2.1). This means that, /l/ is more articulatorily compatible with pharyngealization than /n/ at the tongue root specifically; therefore, it tends to assimilate or coarticulate more than /n/ which tends to consistently resist.

On a separate note, there are a few instances of an overlap between the tongue splines of C in plain context, C in pharyngealized context and C^ʕ (presented in Table 4.2). For instance, the tongue splines of s/_C, s/_C^ʕ and C^ʕ overlapped for 4 speakers. I perceive the consonants as different in the three conditions, but a test should be conducted with naïve listeners. There are two possible scenarios to account for this unexpected overlap. First, it is possible that the pharynx is constricted using ways other than retracting the tongue root to produce the pharyngealization quality. Evidence from the existing literature yields support to this prediction (as explained in §1.3.1.2.2). For example, [E. Al-Tamimi and Heselwood \(2011\)](#) has found that the posterior and lateral walls of the pharynx, along with the tongue root, participate in its cross-sectional

compression by moving them inward. This means that the root retraction would not be the only means of producing pharyngealization; therefore, it could be that the few speakers whose splines overlap in the three conditions constrict the pharynx using the posterior and lateral walls, but without retracting the root. While the ultrasound data used in this dissertation cannot test the role of the posterior/lateral walls in the production of pharyngealization, but other methods such as videofluoroscopy and MRI could. The second scenario that could address the observed overlap is the presence of outliers which might have skewed the results. I inspected the SS-ANOVAs of individual words in each condition and each instance/repetition of these words and they demonstrate that all the target consonants were produced with the same tongue shape and root position.

CHAPTER 5

Physiological Constraints on Coarticulation

5.1 Introduction

Although the human tongue is capable of performing a wide range of complex motions, its hydrostatic properties often induce restrictions on achieving two articulatory gestures concurrently. These antagonistic gestures can be at the same articulator (such as tongue root advancement and retraction) or at different, yet functionally dependent, articulators (such as the blade and dorsum). This gestural antagonism can be a source of complexity in coarticulation. When coarticulation requires performing a gesture that conflicts with the articulatory configuration of the segment subject to coarticulation, different strategies can be used to reconcile the conflict. Reconciliation strategies vary cross- and intra-linguistically.

In this chapter, I report on an ultrasound study investigating what gestures can be potentially incompatible with tongue root retraction and what strategies can be employed to solve the gestural conflict. I use coarticulatory pharyngealization in Arabic as a trigger of tongue root retraction and examine the response to coarticulation from different consonants with varying degrees of gestural antagonism. This antagonism with pharyngealization may arise from two potential lingual gestures. The most severe conflict may stem from two simultaneous antagonistic gestures: advancing the tongue root and raising the tongue body as in producing palatals /j ʃ/ which means applying two forces on the root and body pushing them in the opposite direction of

pharyngealization. A less severe conflict may be associated with tongue dorsum raising as in producing [g x]. Undergoing coarticulatory pharyngealization in this case means retracting the root and raising the dorsum and body concurrently which, in turn, means forcing the tongue beyond what its volume can permit. In this chapter, I investigate the production of the two types of gestural antagonism by examining the articulation of words containing velar or palatal consonants when occurring in a pharyngealized context.

This chapter is organized as follows. I first outline the research question, hypotheses and methodology. The Results section starts with an overview of the articulation of phonemic pharyngealization followed by a description of the effects of assimilatory pharyngealization on all types of tested consonants, constrained and unconstrained. The Discussion presents an interpretation of the findings on what gestures are antagonistic to coarticulatory pharyngealization and what strategies are used to resolve this conflict.

5.2 Research Question and Hypotheses

A few potential gestural conflicts with pharyngealization originate from the physiological constraints on the tongue movement, most of them having to do with the inter-dependence between lingual articulators, as laid out in detail in Chapter 2. The key question that is addressed in this chapter is: what are the reconciliation strategies that speakers use to solve this type of conflict in coarticulation? Given the non-automaticity of articulatory conflict reconciliation as discussed in §2.3.2.2, two hypotheses were formulated to describe possible scenarios of reacting to conflict emerging from coarticulatory pharyngealization.

5.2.1 Compromise hypothesis. A possible reconciliation strategy is a compromise between antagonistic gestures resulting in an articulatory reduction of both gestures. The tongue should react within the limits of its mechanical capabilities (i.e. without violating its volume-

preservation property) but, at the same time, it should allow coarticulation to take place. That is, in producing a gesturally antagonistic consonant (e.g. /x g j ʃ/), the root should be retracted to exhibit coarticulatory variability but to a lesser extent than that of a phonemically pharyngealized consonant and the dorsum should be raised but to lesser extent than in the canonical form. This reduction in phonemic dorsal raising, however, should not lead to the loss of the consonant's identity; the place of articulation of the consonant should be comparable to that of its non-coarticulated counterpart. In this sense, these consonants should be less sensitive to coarticulation than consonants with no such articulatory antagonism, for example, labials and dentals.

[Recasens et al. \(1997\)](#) have shown that there is a positive correlation between articulation constraints (or coarticulation resistance) and dominance and that consonants that are highly constrained (and often resistant to coarticulation) can actually exert a coarticulatory effect on other neighboring segments. This two-way effect predicts that the constrained consonants (e.g. [x g j ʃ]) should not only restrain coarticulatory root retraction during their articulation (to some extent) but should also exert their own coarticulatory effect on C^s by raising the dorsum to some extent and pulling the root forward away from its canonical retracted position. A potential consequence of this compromise would be the raising the intervening lax low vowel. For example, /ʃatʰarha:/ 'he cut her into two halves' could surface as [ʃetʰarha:], a possibility that has never been attested.

5.2.2 Resistance hypothesis. Another possible scenario is resistance to coarticulatory root retraction in order to maintain the phonemic, canonical dorsal raising of the target consonant. Two types of gestural conflicts arise from the root retraction in pharyngealization: a same-articulator conflict induced by the root advancement of the target consonants /ʃ j/, and an inter-articulator conflict by the dorsal raising of the target consonants /x g/. This hypothesis states that same-

articulator conflict yields more extreme resistance to coarticulatory root retraction than inter-articulator conflict while sustaining the phonemic dorsal raising in the latter case.

There is, in fact, a third scenario where an excrescent schwa is inserted between the conflicting gestures, analogous to English [ə] between advanced [i] and retracted liquid ([Gick & Wilson, 2006](#)) discussed in Chapter 2. However, this is likely to occur in cases where the conflicting gestures are immediately adjacent to each other. In the present study, however, the phones with conflicting gestures are separated by an inter-consonantal vowel /a/ and, therefore, this hypothesis of schwa insertion is ruled out.

5.3 Methodology

The articulation experiment was designed to examine the susceptibility to regressive coarticulation in three sets of consonants with varying degrees of lingual constraints and gestural antagonism. The consonants can be either unconstrained lingually, or moderately or highly antagonistic to pharyngealization. The experiment followed the same design explained in Chapter 3-4 and the data were obtained in the same recording session as the experiment in Chapter 3-4 but with a different wordlist. Mid-sagittal ultrasound images of the tongue were collected while 15 native speakers performed a sentence reading task. The selected ultrasound images were traced manually using Palatoglossatron and SS-ANOVAs were fitted to the tongue traces.

The wordlist (presented in Table 5.1) consisted of disyllabic and multisyllabic words containing the sequence C₁aC₂ where C₁ is the target consonant subject to coarticulation and C₂ is the coarticulation source consonant. The words comprised (near)-minimal pairs contrasting in pharyngealization where the source C₂ was either pharyngealized /tʰ sʰ/ or plain /t s/. The target consonants, which always preceded the source consonants in the experiment, varied in their place of articulation (henceforth, ‘target place’): either labial [b f], palatal [ʃ j] or velar [x ɣ]. The manner

of articulation of the target consonants also varied. For each place of articulation, a stop (e.g. [b] and [g]) and a fricative (e.g. [f], [x] and [ʃ]) were selected, except for the palatal place that had a glide [j] instead of a stop because Arabic does not have a palatal stop such as /ʃ c/.

There were two pairs of each combination of C₁ and C₂, except for the combination of the palatal /ʃ/ and the source /s sʕ/ because of the lack of words containing the sequence /ʃasʕ/ in the language. This resulted in a total of 44 words, comprising 16 labial, 16 velar and 12 palatal words. The results presented in this study were based on the analysis of 15,038 ultrasound images: 5,712 images for words containing labials, 4,291 images for palatals and 5,035 images for velars.

Table 5.1

Complete list of words used to address gestural antagonism showing stress with IPA diacritic and syllable structure

Target place	Target	Source	Example	Gloss
Labial	b	t	ba.ta.'la:t	petals-FEM
		tʕ	ba.tʕa.'la:t	heroines-FEM
		t	ba:t.'ha:	slept it/her
		tʕ	ba:tʕ.'ha:	blew her up
		s	ba:s.'ha:	he kissed her
		sʕ	ba:sʕ.'ha:	her bus
		s	ba.sam.'ha:	he smiled her
		sʕ	ba.sʕam.'ha:	he took her finger print
	f	t	fa:t.'na:t	attractive-FEM-PL
		tʕ	fa:tʕ.'na:t	smart-FEM-PL

Target place	Target	Source	Example	Gloss
		t	fa:t.'ra:t	cold-FEM-PL
		t ^ʰ	fa:t ^ʰ .'ra:t	eating breakfast-FEM-PL
		s	fa.sad.'ha:	he ruined her
		s ^ʰ	fa.s ^ʰ al.'ha:	he fired her
		s	'ra.fas	he kicked
		s ^ʰ	'ga.fas ^ʰ	cage
Palatal	j	t	jat.le.'fu:n	they are destroying
		t ^ʰ	jat ^ʰ .le.'bu:n	they are asking
		t	ja.tu:.'bu:n	they are repenting
		t ^ʰ	ja.t ^ʰ u:.'fu:n	they are walking around
		s	ja.sed.'ha:	he closes her
		s ^ʰ	ja.s ^ʰ ed.'ha:	he stops her
		s	jas.ref.'ha:	he is ladling her
		s ^ʰ	jas ^ʰ .ref.'ha:	he is withdrawing her
	ʃ	t	ʃa.tar.'ha:	he cut her into pieces
		t ^ʰ	ʃa.t ^ʰ ar.'ha:	he cuts her into two halves
		t	ʃa:t.'ma:t	insulting-FEM-PL
		t ^ʰ	ʃa:t ^ʰ .'ba:t	omitting-FEM-PL
		s, s ^ʰ	(Not found)	(Not found)
Velar	g	t	ga:t.'la:t	killer-FEM-PL
		t ^ʰ	ga:t ^ʰ .'fa:t	pickers-FEM-PL (of fruit)
		t	'la.ga.tah	she found him/it

Target place	Target	Source	Example	Gloss
		tʰ	'la.ga.tʰah	he caught him
		s	ga:s.'ha:	he measured her
		sʰ	ga:sʰ.'ha:	he narrated/cut her
		s	ga.sam.'ha:	he divided her
		sʰ	ga.sʰər.'ha:	her palace
	x	t	xa.tam.'ha:	he stamped her
		tʰ	xa.tʰab.'ha:	he proposed to her
		t	'sa.xat	to be generous-FEM
		tʰ	'sa.xatʰ	to be angry
		s	xa.sar.'ha:	he lost her
		sʰ	xa.sʰam.'ha:	he discounted her
		s	xas.'ha:	her lettuce
		sʰ	xasʰ.'ha:	he made private

5.4 Results

For every comparison in this section, a representative speaker's SS-ANOVA (at 50% of the phone/closure) is presented along with a summative plot that summarizes the SS-ANOVA findings of all individual speakers and a third plot to showing the difference between all frames. As in the previous two chapters, the speaker that is chosen represents a sub-sample of speakers in terms of the overall tongue shape, the direction of differences between tongue fits and approximates the size of a difference between tongue fits.

Before delving into coarticulation and gestural antagonism, I will demonstrate that the articulatory configuration of phonemic pharyngealization in source consonants is the same in this part of the experiment as in previous chapters. Unlike the plain consonants /t s/, their phonemically pharyngealized counterparts /t^ʕ s^ʕ/ occurring after labials, palatals and velars were characterized by a retraction of the tongue root and a flattening and lowering of the tongue body (Figure 5.1). This tongue configuration was also observed for all speakers and at all selected time points: at the release of /t^ʕ/ and at 75% and 25% of /s^ʕ/ (Figure 5.2).

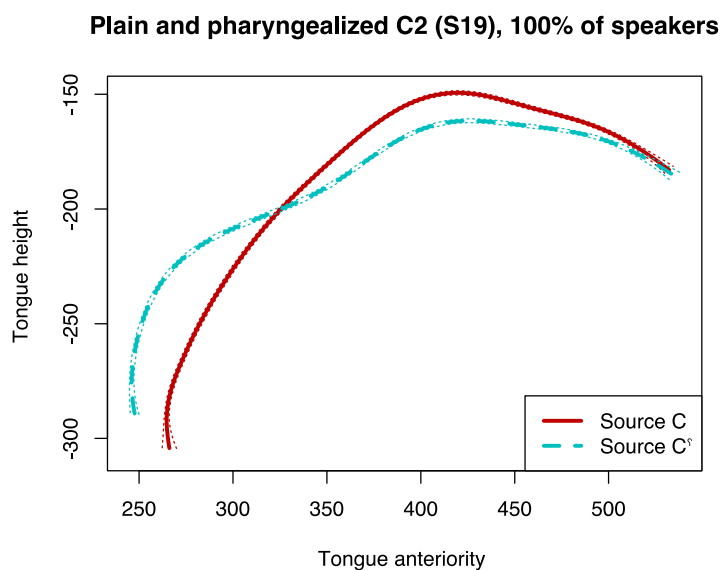


Figure 5.1. SS-ANOVA fit at mid-phone/closure and 95% CI showing plain and pharyngealized C2.

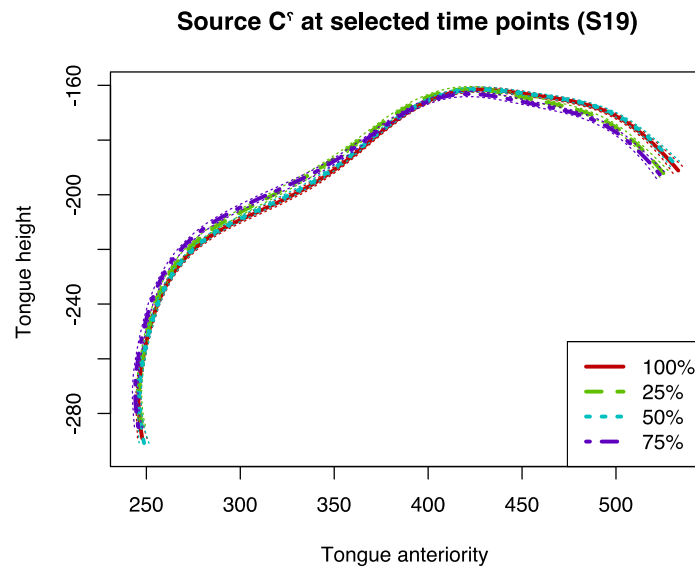


Figure 5.2. SS-ANOVA fit and 95% CI showing C^s at all selected time points.

The fricative /s^s/ and the stop /t^s/ had the same tongue configuration and the same degree of root retraction (Figure 5.3), except for two speakers who showed a negligibly greater retraction in /s^s/ than /t^s/.

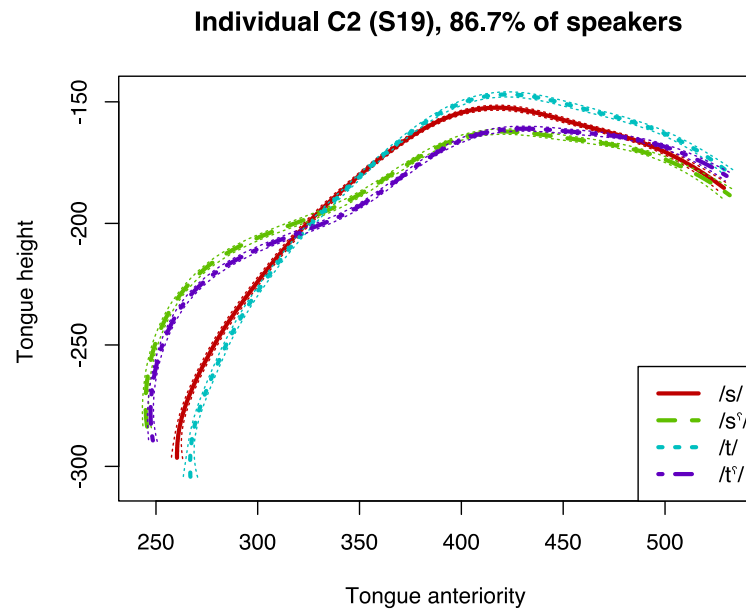


Figure 5.3. SS-ANOVA fit at mid-phone/closure and 95% CI showing all C2s /t t^s s s^s/.

In order to determine whether the palatals and velars induce progressive coarticulation on the following C^s , the extent of tongue root retraction in source C^s was examined across the three target places, labials, palatals and velars (Figure 5.4). It was found that /t^s s^s/ were produced with the same extent of phonemic root retraction in all contexts among all speakers except 3: one of them produced C^s in a labial context with marginally more root retraction, and two produced C^s in a palatal context with a slightly less root retraction. The difference in the root position in these three cases was trivial.

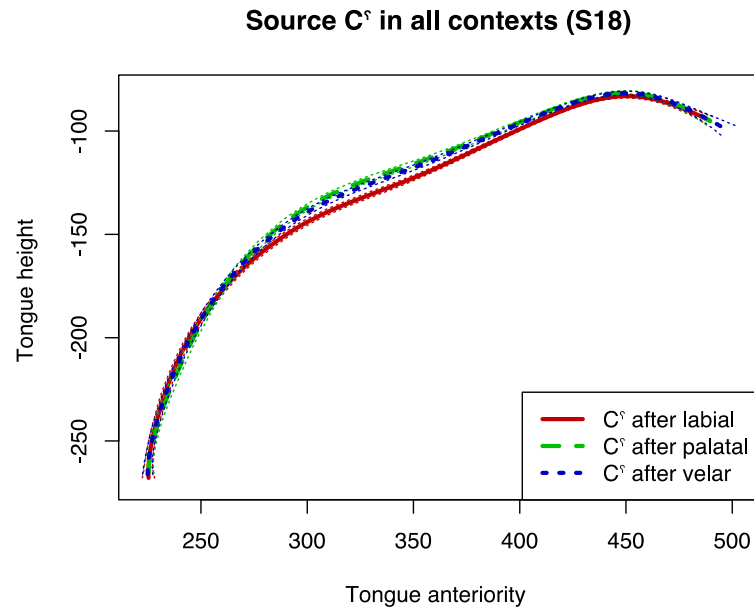


Figure 5.4. SS-ANOVA fit at mid-phone/closure and 95% CI showing C^s preceded by labials, palatals and velars.

5.4.1 Labials. Bilabial [b] and labiodental [f] were used in this study as a control since they do not involve any lingual articulatory movement that could interfere with pharyngealization. This was evident in the ultrasound images of labials in plain context e.g. /basamha:/ ‘he smiled’ (Figure 5.5, left) which shows that the tongue root was in a neutral position and that the dorsum

was not as raised as in some lingual consonants such as the velars. This tongue shape was completely different from that of labials preceding source C^s e.g. /bas^samha:/ ‘he took a fingerprint’ where the tongue root was retracted and the tongue body was lowered and flattened (Figure 5.5, right), a shape qualitatively similar to that of phonemically pharyngealized consonants.

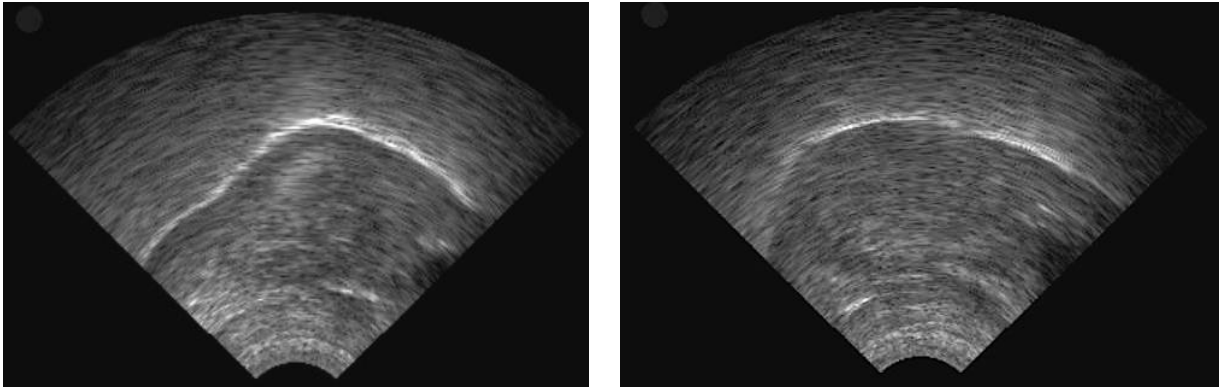


Figure 5.5. Ultrasound images from speaker S17 at mid-closure of labial [b] in plain context (left) and pharyngealized context (right) from the same pair /basamha: - bas^samha:/.

SS-ANOVA findings across all speakers show that labials before source C^s had more tongue root retraction than labials before source C for the majority of speakers (73.3% of 11/15 speakers). This retraction was less extreme than the retraction in phonemic pharyngealization (i.e. in producing source C^s) for some of them (53.3% or 8/15) and as extreme as source C^s for others (20% or 3/15), both of which are presented in Figure 5.6 and 5.7, respectively.

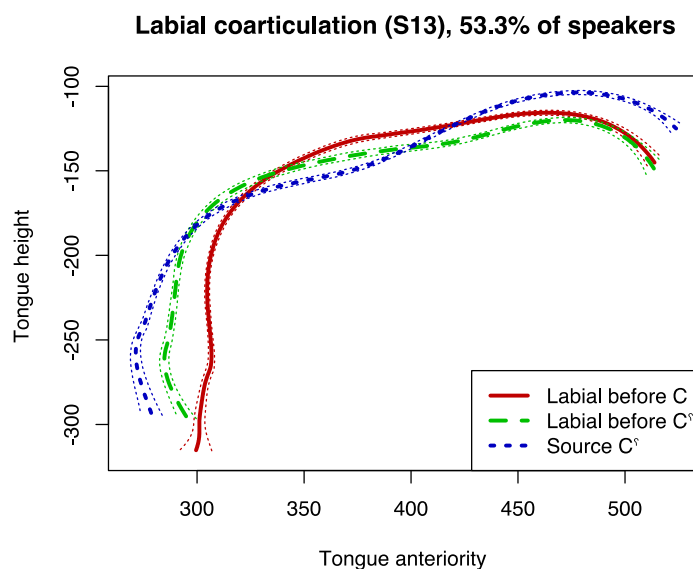


Figure 5.6. SS-ANOVA fit at mid-phone and 95% CI showing coarticulation in labials.

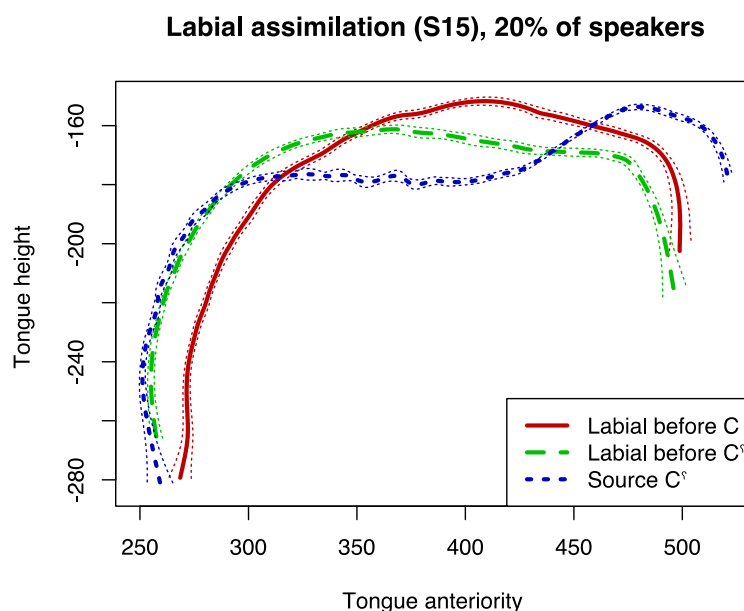


Figure 5.7. SS-ANOVA fit at mid-phone and 95% CI showing assimilation in labials.

Within-speaker comparisons of labials in plain versus pharyngealized context and versus source C^s are summarized in the scatterplot in Figure 5.8. As in the previous two chapters, the x-

axis represents the minimum difference (per speaker) between SS-ANOVA fits of the two labial conditions (i.e. fit of labials in a pharyngealized context - fit of labials in a plain context) at the tongue root, an angle from horizontal to 45° ($0, \pi/4$). The y-axis refers to the difference between the two labial conditions at the tongue body area: from 75° to 105° . As illustrated in Figure 5.8 (left) where the root difference is > 0 , there was a backward displacement of the tongue root in articulating labials in a pharyngealized context compared to labials in plain context across all speakers. The displacement ranged between 5 and 30mm. Speakers also varied in the tongue body height in the two labial conditions, with the tongue body being lower in pharyngealized labials than plain ones. The degree of root retraction in coarticulatory pharyngealization is not the same as in phonemic pharyngealization for many speakers, as illustrated in Figure 5.8 (right). For the majority of speakers, the tongue root was more retracted in producing the source C^s than in producing labials in pharyngealized context as the root difference was < 0 for many speakers.

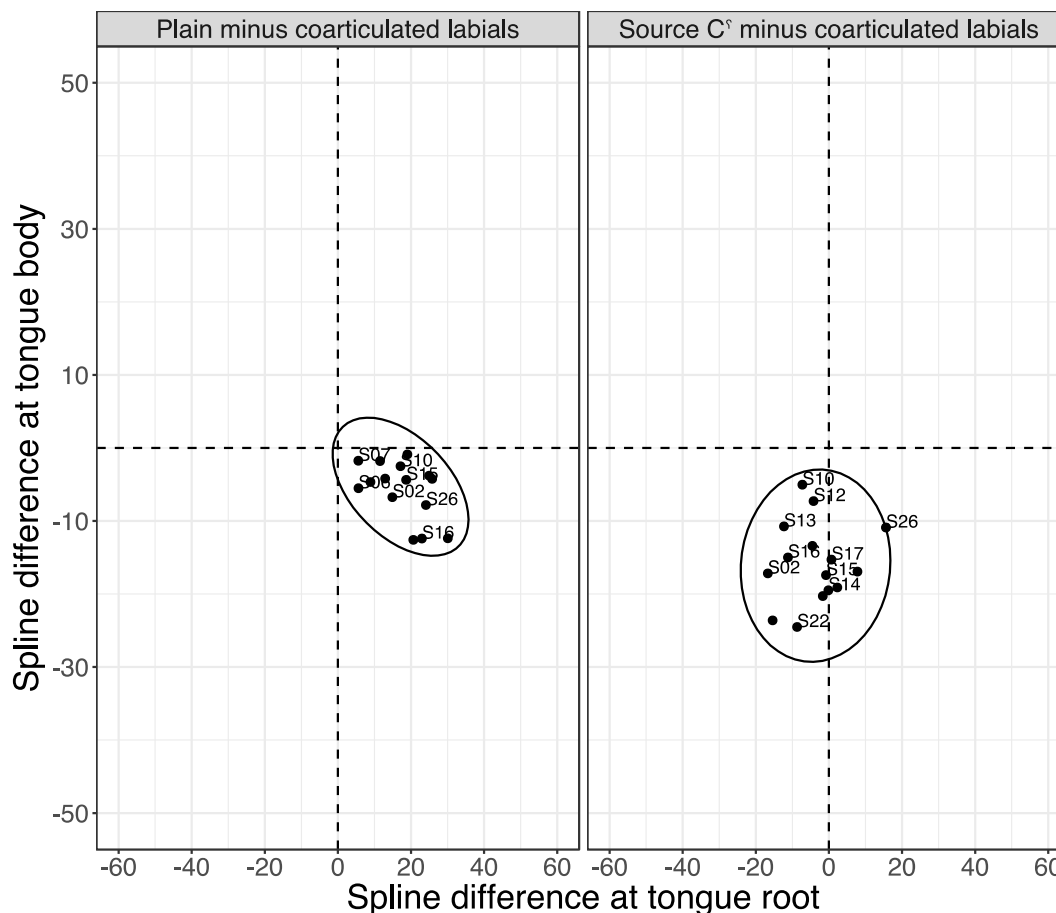


Figure 5.8. The difference between SS-ANOVA fits of coarticulated labials vs plain ones (left) and vs distinctive C^s (right) for all speakers. Positive root difference in the left means that the coarticulated labials are more retracted than plain ones.

Similar to the previous chapter, the classification of the effect was based on the results of the ratio and ANOVA of the radius at the tongue root area. The root retraction ratio of [b f] in a plain context was lower than the ratio of [b f] in a pharyngealized context, and either lower or as high as the retraction ratio of the source C^s depending on whether the effect is gradient or categorical. For the 80% of speakers who demonstrated an assimilatory effect in labials, the root radius of coarticulated labials was largely different from labials in a plain context with $F > 30$. For example, the difference between the two labial conditions for the speaker presented in Figure 5.9

was $F(1,71) = 119.97, p < 0.001$, which indicates the presence of a coarticulatory retraction in a pharyngealized context. The root radius of coarticulated labials by this speaker who demonstrated a gradient effect was also different from the root radius of the source C^{c} : $F(1,114) = 211.12, p < 0.001$, which indicates that the coarticulatory retraction was less extreme than in the source C^{c} . The F-value of the difference between assimilated labials and the source C^{c} was < 30 for the speakers who assimilated labials categorically.

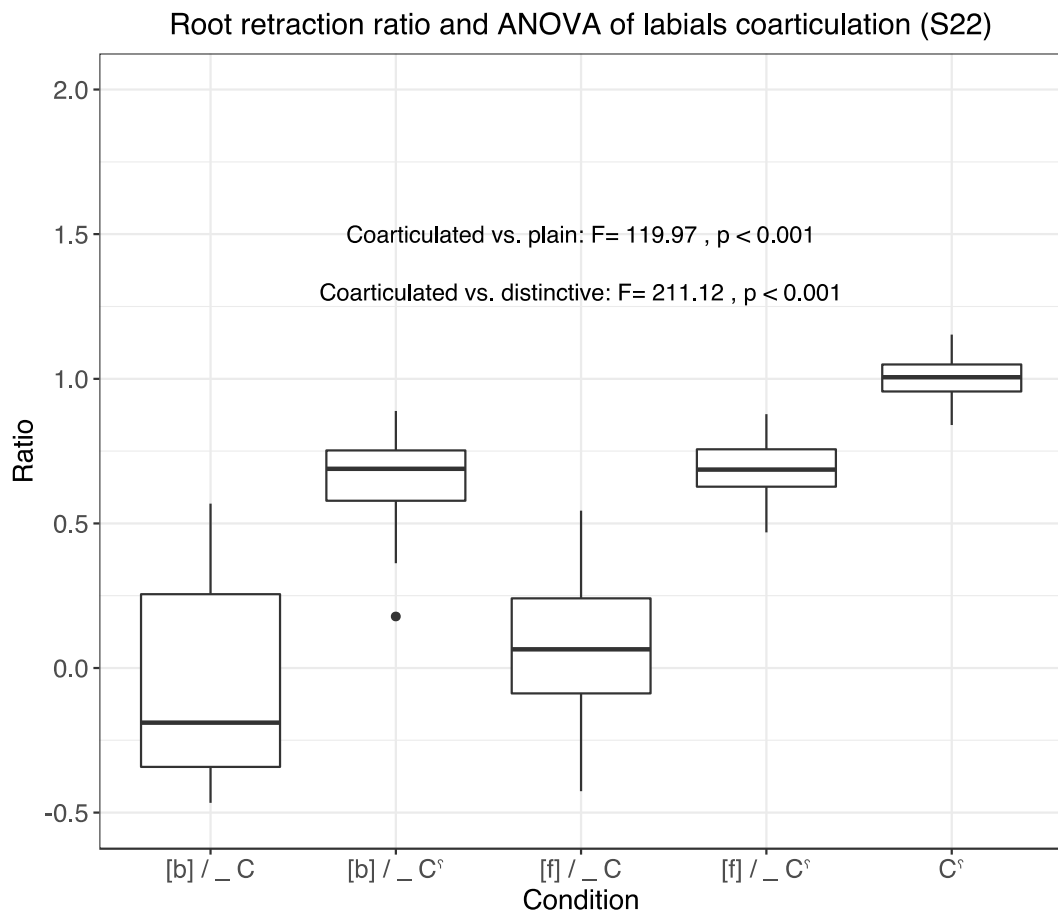


Figure 5.9. ANOVA and ratio of mean root radius showing coarticulation in labials.

The rationale for combining the results of the two labials together was that both showed the same degree of coarticulatory tongue root retraction at 50% of the phone/closure (Figure 5.10).

This was due to the fact that the bilabial [b] and the labiodental [f] are equally unconstrained at the tongue body, dorsum and root.

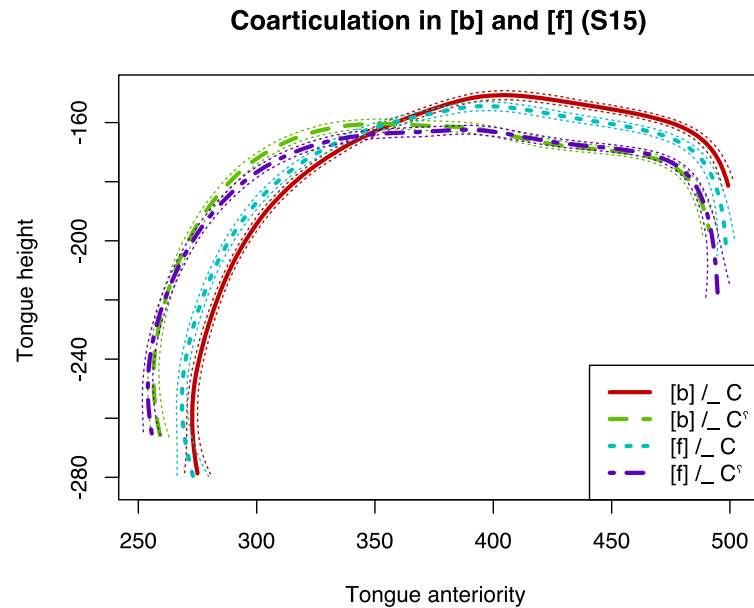


Figure 5.10. SS-ANOVA fit at mid-phone and 95% CI showing coarticulation in [b] vs [f].

This tongue shape at 50% of the phone/closure is not different from the other time points selected from the labial fricatives and stops, as shown in Figure 5.11, suggesting that magnitude of coarticulatory pharyngealization in labials is persistent throughout the entire consonant and that it is realized articulatorily as early as the mid of the holding phase of [b] and the first 25% of [f] friction. The only tongue fit that was consistently more retracted than the other fits was the one at 75% of the phone, suggesting that the retraction becomes greater towards the end of the phone that is close to the source C^s, but this difference was small as shown in Figure 5.11.

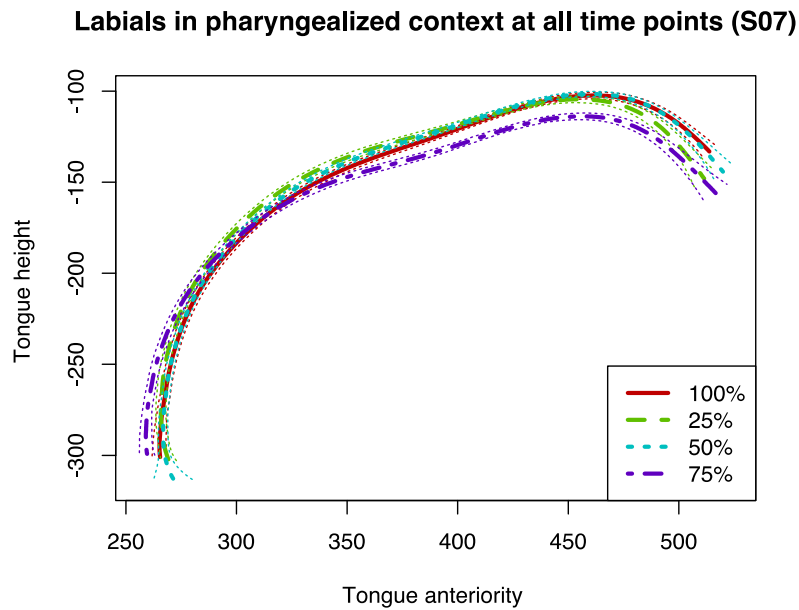


Figure 5.11. SS-ANOVA fit of labials in pharyngealized context at all selected time points.

5.4.2 Palatals. Palatal [ʃ j] are produced with an advanced tongue root and therefore present a same-articulator conflict with coarticulatory root retraction in pharyngealized words. This conflict was expected to result in either a reduction in phonemic root retraction to allow coarticulation (compromise hypothesis) or a high resistance to coarticulation to maintain the phonemic root posture (resistance hypothesis).

Ultrasound data showed root advancement and dorsal raising in articulating the palatals in both plain and pharyngealized contexts, indicating that they were produced with similar tongue configuration. This is illustrated in Figure 5.12 where the tongue spline of both palatal [ʃ j] preceding C^s overlapped with the spline of the same consonants preceding plain C.

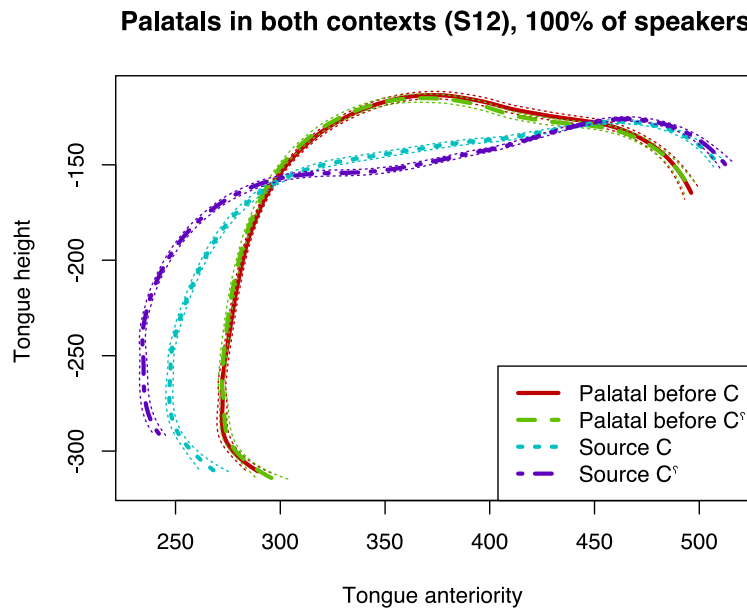


Figure 5.12. SS-ANOVA fit at mid-phone and 95% CI showing palatals [f j] in both contexts.

The two palatals differed from each other (for some speakers) in the anteriority of the dorsal constriction where [j] required pre-dorsal raising and [f] required post-dorsal raising, as in Figure 5.13. Therefore, the two consonants were analyzed separately in this study. Despite this difference, the tongue configuration of both consonants did not change before the source C^s - that is, both were completely resistant to coarticulation. This finding eliminates the compromise hypothesis (which predicts lowering the pre-dorsum and slightly reducing the root advancement) and supports the resistance hypothesis because the phonemic gestures (i.e. pre-dorsal raising and root advancement) were not reduced in the pharyngealized context but were rather maintained.

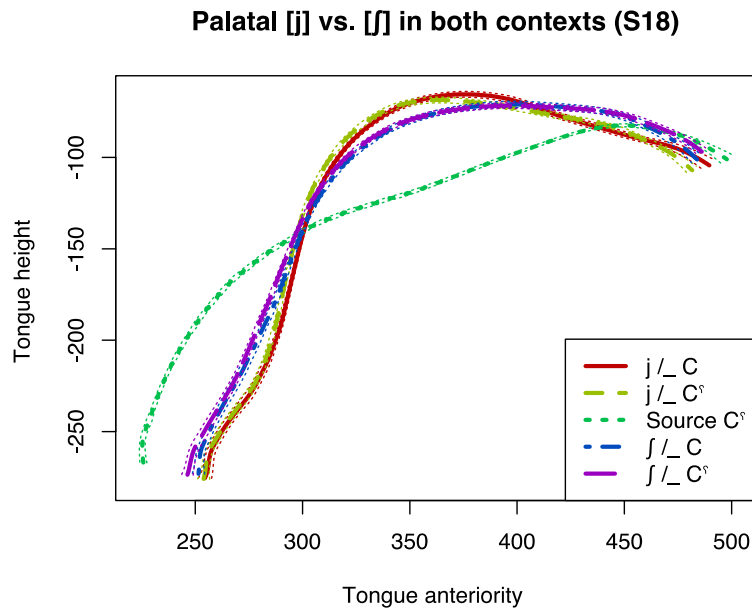


Figure 5.13. SS-ANOVA fit at mid-phone and 95% CI showing palatal [j] vs. [ɟ] in both contexts.

Resistance of palatals to pharyngealization was true for most speakers as shown in Figure 5.14 which compares palatals preceding $C^{\text{ɹ}}$ to the two reference points: palatals in a plain context and the source $C^{\text{ɹ}}$. The differences were measured at tongue root and body of the SS-ANOVA tongue fits of all individual speakers. It is evident from the left scatterplot that, in reference to palatals in plain context, the tongue root backward displacement in producing both [j] and [ɟ] in the pharyngealized context was marginal. There was even less tongue body lowering in the pharyngealized context (i.e. less than 10 mm) by all speakers except one. As in Figure 5.14 (right), a comparison of the tongue root position in producing the source $C^{\text{ɹ}}$ versus palatals occurring in a pharyngealized context showed a large forward displacement of the tongue root in producing the two palatals ranging between 10 and 60 mm and a large upward displacement of the tongue body (~20-50 mm). The scatterplot also shows that the fricative [ɟ] had a higher tongue body than the approximant [j] among most speakers. These scatterplots illustrate that, across all speakers, 1)

resistant palatals exhibited marginal or no articulatory change when occurring before the source C^s ; 2) the tongue root displacement within the pharyngeal space was remarkably different from the source C^s ; and 3) the two palatal consonants [j] and [ɟ] shared the same resistance pattern although they differed in tongue body height.

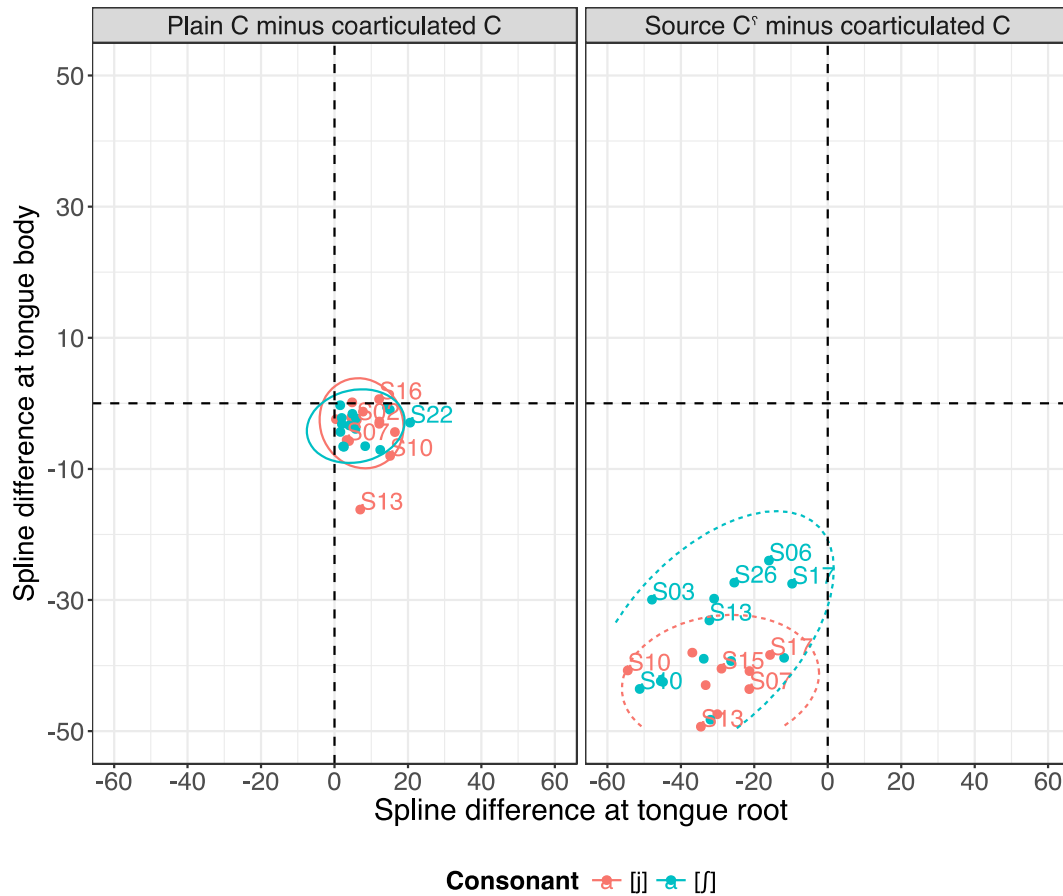


Figure 5.14. The difference between SS-ANOVA fits for individual palatals [j] [ɟ] in pharyngealized vs plain context (left) and vs source C^s (right) for all speakers. Negative root and body difference in the right means that the palatals had less retracted root and higher body than source C^s .

The root retraction ratio of [j] and [ɟ] preceding a plain C was ≤ 0 and did not increase substantially when both preceded a pharyngealized source C^s (Figure 5.15). What this means is that the palatal approximant [j] was produced with a more advanced root than the fricative [ɟ] and

that the advancement in both consonants was not affected by the upcoming pharyngealized consonant. ANOVAs revealed that the difference of the root radius between [j] in a plain context and in a pharyngealized context and between [ʃ] in a plain context and in a pharyngealized context was insignificant and $F < 30$ for all speakers. This applied, for example, for the speaker whose SS-ANOVA and retraction ratio are presented in Figure 5.13 and 5.15 ([j]: $F(1,37) = 0.18, p > 0.001$; [ʃ]: $F(1,18) = 0.79, p > 0.001$). The position of the tongue root in producing both palatals was completely distinct from the source C^s , as indicated by the ANOVA of the root radius for all speakers. For example, the radius difference between [j] and C^s for the speaker whose results appear in Figure 5.13 and 5.15 was $F(1,120) = 1269.5, p < 0.001$ and [ʃ] vs. C^s was $F(1,110) = 354.56, p < 0.001$.

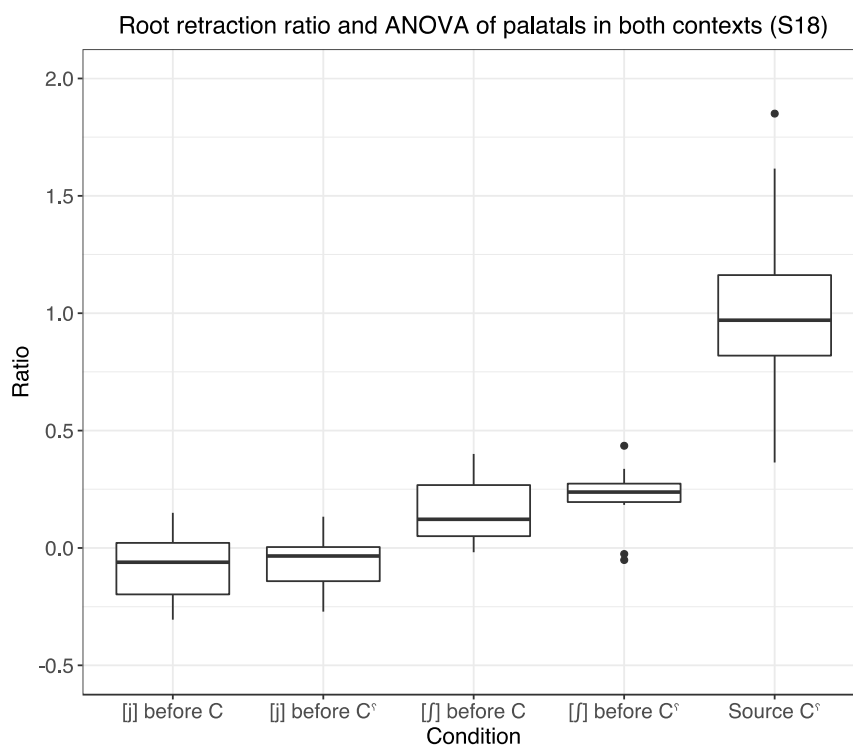


Figure 5.15. ANOVA and ratio of mean root radius of palatal [j] and [ʃ] in both contexts.

Lastly, the root advancement of palatals in pharyngealized words was consistent across all other selected time points. This was true for both [j] and [ʃ] (as in Figure 5.16 and 5.17, respectively).

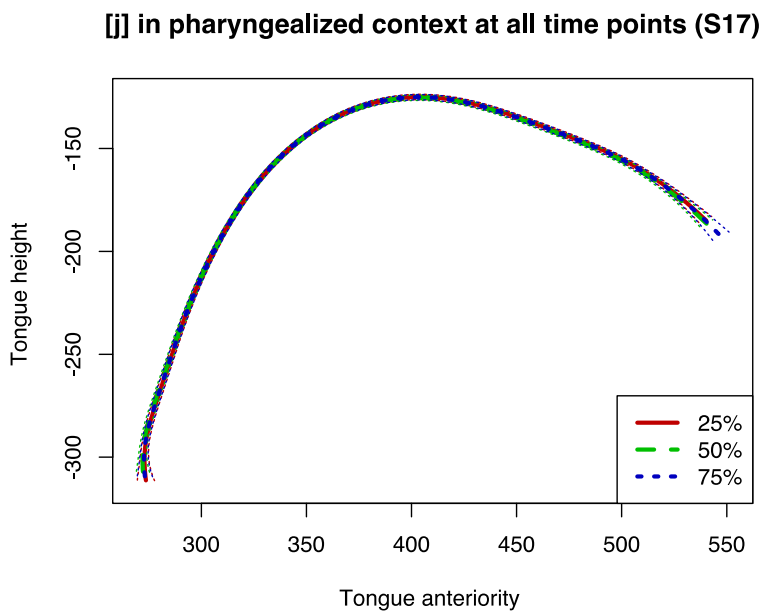


Figure 5.16. SS-ANOVA fit showing [j] in pharyngealized context at all selected time points.

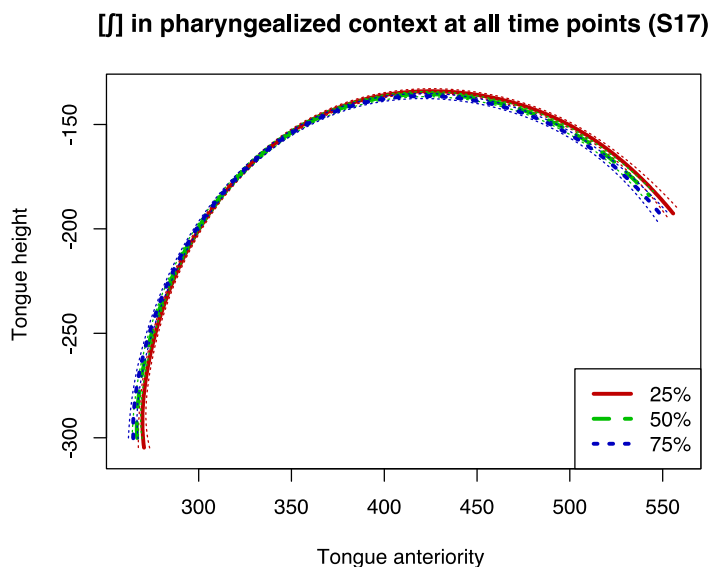


Figure 5.17. SS-ANOVA fit showing [ʃ] in pharyngealized context at all selected time points.

5.4.3 Velars. The two consonants [x g], produced with dorsal elevation, were predicted to be in an inter-articulator conflict with the coarticulatory tongue root retraction induced by pharyngealization. In fact, since it was found in Chapter 3 that pharyngealization was not only produced with a tongue root retraction but also with a mechanical tongue body lowering, the conflict between velars and pharyngealization is more complex than assumed in the formulation of the hypotheses. There is not only a gestural antagonism between the dorsal raising of velars and the root retraction of pharyngealization due to the tongue's volume preservation, but also a gestural antagonism with the mechanical tongue body lowering caused by pharyngealization.

The ultrasound data revealed that unlike the two palatals which were articulated with a similar tongue shape, the two velars were different from each other in their articulatory configuration. Consistently among all speakers, the stop [g] in the plain context was produced with a (pre)-dorsal constriction against the palate and tongue root advancement, a configuration similar to that of the palatal consonants except that the palatal contact in [g] was tighter and the pre-dorsum was higher. An example lingual shape in producing [g] (in a plain context) relative to the source C₂ and C₂^ϕ is presented in Figure 5.18. The articulation of the other consonant [x] was not expected: the dorsum was raised and retracted towards the velum and in these cases of dorsum retraction, the tongue root was not retracted. The involvement of the tongue root was variable across speakers: the tongue root in [x] was retracted to the same extent as in C₂^ϕ by 53.3% of speakers (Figure 5.18) and not retracted by 46.7% of speakers (Figure 5.19). This tongue configuration suggests that this consonant is not produced as velar in this dialect, but is rather uvular and should therefore be transcribed as [χ]. Given this, in what follows, I will use the symbol [χ] instead of [x] and 'uvular' instead of 'velar' to refer to this consonant. Regardless of this speaker variation in the pharyngeal constriction, the dorsal height was consistent across all speakers in [χ] which was consistently

articulated with a lower dorsum than [g]. The tongue in [χ] was lower than [g] to allow air to pass through the constriction and produce the friction noise, unlike the stop [g] which required a higher dorsum to form a complete stop closure.

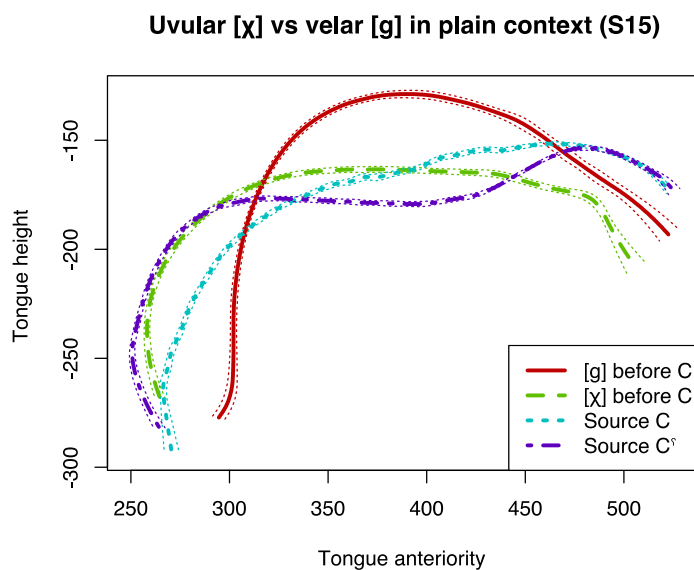


Figure 5.18. SS-ANOVA fit at mid-phone and 95% CI of [χ] vs. [g] in plain context, with [χ] retraction.

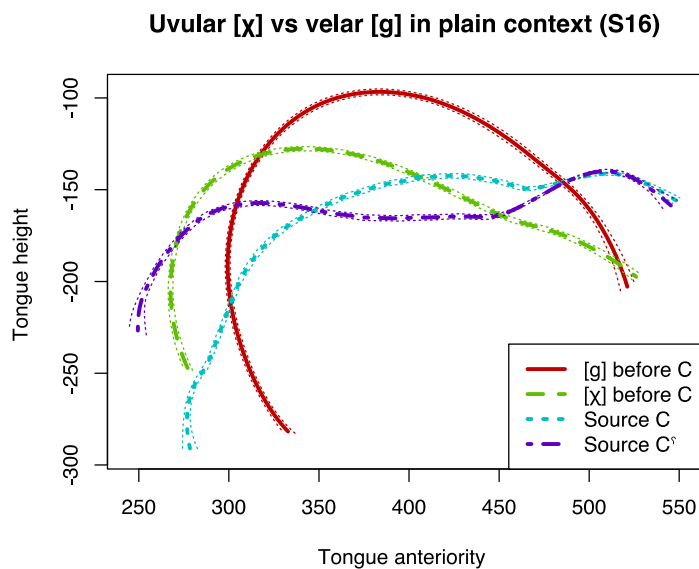


Figure 5.19. SS-ANOVA fit at mid-phone and 95% CI of [χ] vs. [g] in plain context with no [χ] retraction.

If resistance is the strategy used to solve articulatory conflict with pharyngealization and if it is correlated with the degree of dorsal raising, velar [g] should be more resistant to pharyngealization than the uvular [χ] because it has a high dorsum and, thus, greater articulatory conflict with pharyngealization than [χ]. Another gestural antagonism that should cause more resistance in [g] than [χ] is the position of the tongue root: it is compatible in [χ] but antagonistic in [g]. Also, if resistance is positively correlated with articulatory demands and degree of constriction, the velar [g] should be the most resistant consonant in the study because the dorsum reached the highest position in producing it compared to the other target consonants. Due to these articulatory differences between [g] and [χ], I present the results of both separately.

5.4.3.1 Resistance in [g]. Ultrasound data of the velar [g] in the pharyngealized context showed that the degree of palatal contact and root advancement did not change when it occurred before C^ɣ. As illustrated in Figure 5.20, in many speakers, the tongue fit of [g] in the plain context was close to the fit of [g] in the pharyngealized context.

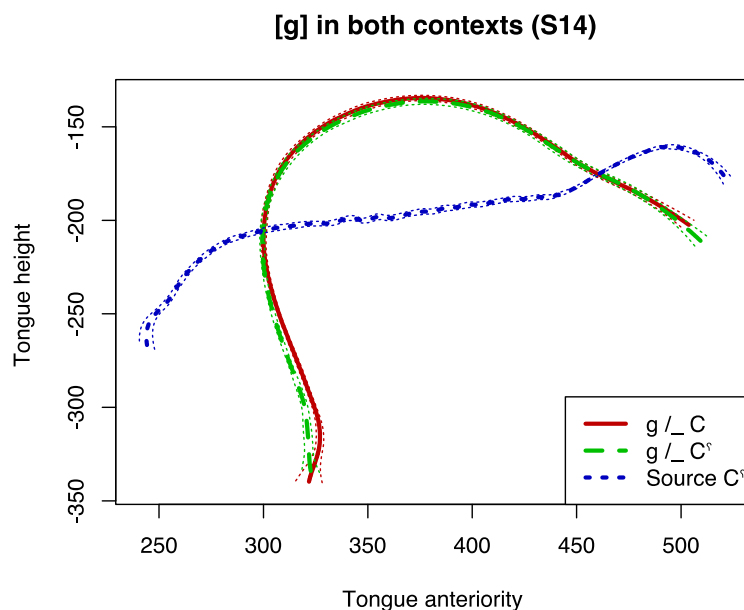


Figure 5.20. SS-ANOVA fit at mid-closure and 95% CI showing resistance velar [g].

A summary of all speakers' SS-ANOVA fits of [g] at the tongue root and body in a pharyngealized context is presented in Figure 5.21. As shown in the left scatterplot, the position of the tongue root and body while producing [g] in a pharyngealized context was close to its position in a plain context. This means that this consonant was resistant to coarticulation by all speakers, with no (or extremely marginal) reduction in root advancement and tongue body raising. Evidently, measurement of the tongue body and root position of [g], in reference to the source C^s , showed a remarkable forward movement of the tongue root and upward movement of the tongue body (Figure 5. 21 right).

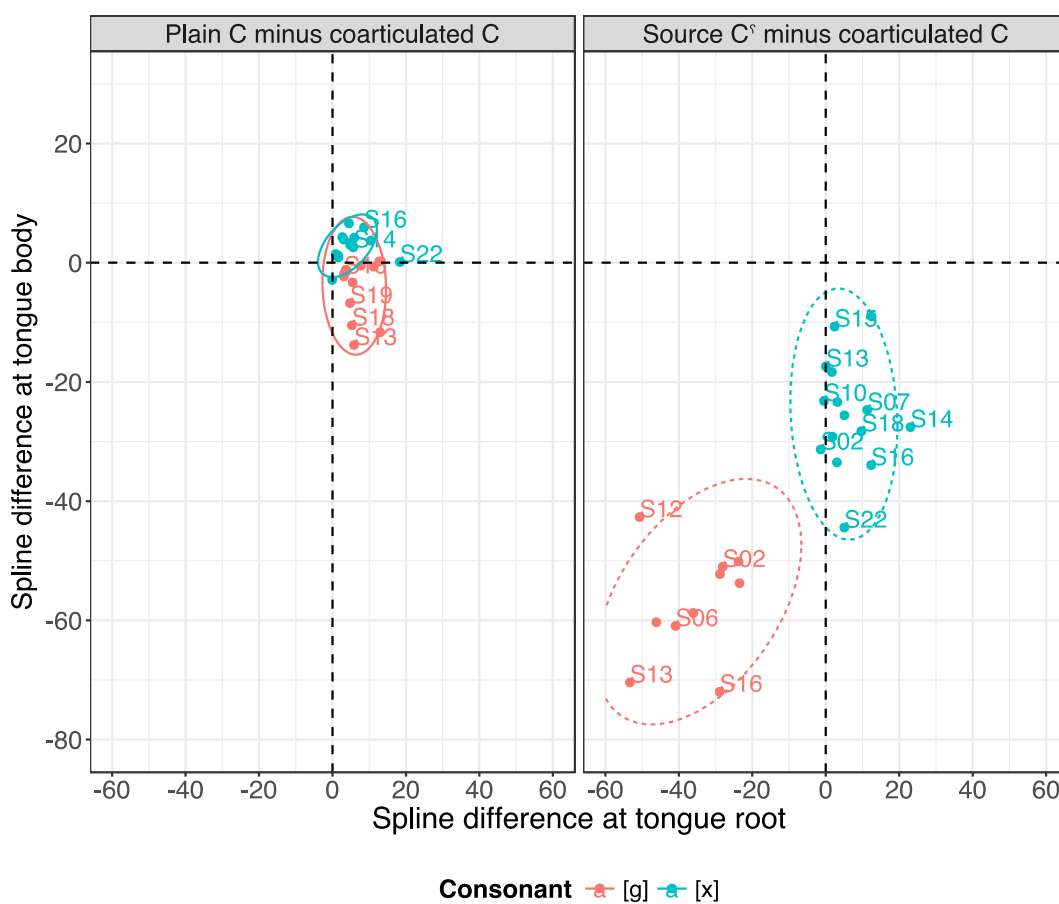


Figure 5.21. The difference between SS-ANOVA fits for individual velars [χ g] in pharyngealized vs plain context (left) and vs source C^s (right) for all speakers. Negative root and body difference in the right means that the velar had less retracted root and higher body than source C^s .

Consistently among all speakers, the root retraction of [g] in the plain context was very close to the ratio of the same consonant in the pharyngealized context (both approximated 0), and ANOVAs on the radius in these two conditions yielded small insignificant differences. For example, as in Figure 5.22, the root radius while producing [g]/_ C^ʕ by speaker S14 (whose SS-ANOVA is presented above) was not sizeably different from the radius of [g]/_ C: $F(1,36) = 5.81$, $p > 0.001$. Tongue root position of [g]/_ C^ʕ was completely distinct from its position in C^ʕ as indicated by a large significant difference between them in terms of their root radius: $F(1,126) = 4008.93$, $p < 0.001$. This large difference reflects two opposite displacements of the tongue root: retraction in C^ʕ versus advancement in [g] although the latter occurred in a pharyngealized context. This root advancement and resistance of [g] was not only observed at 50% of the stop closure but also at the release as presented in Figure 5.23.

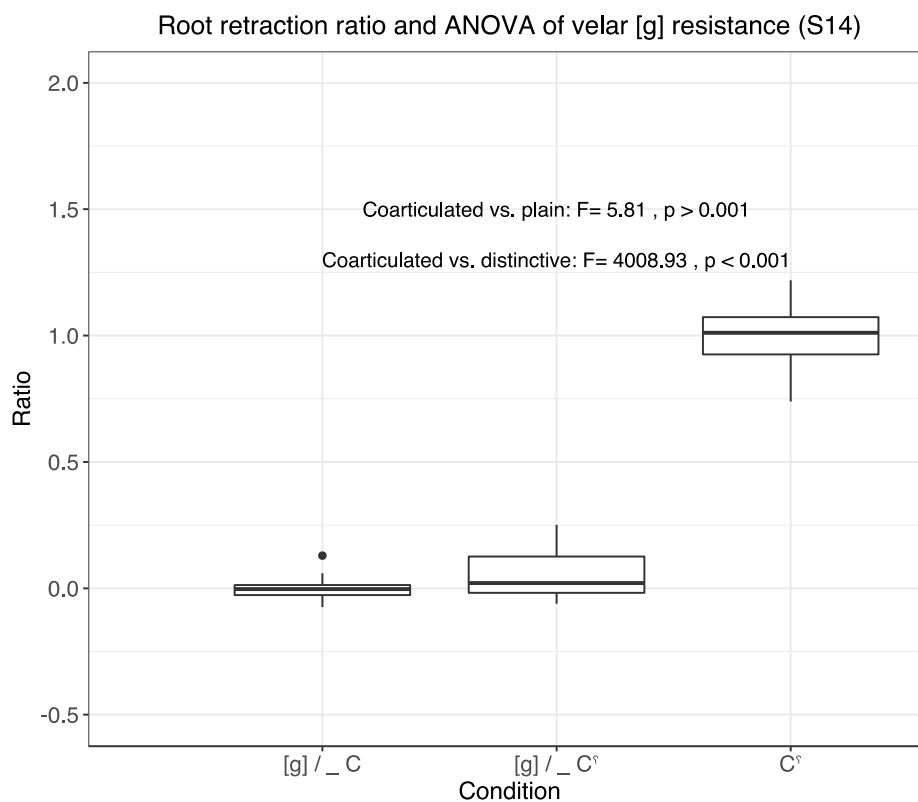


Figure 5.22. ANOVA and ratio of mean root radius showing velar [g] in both contexts.

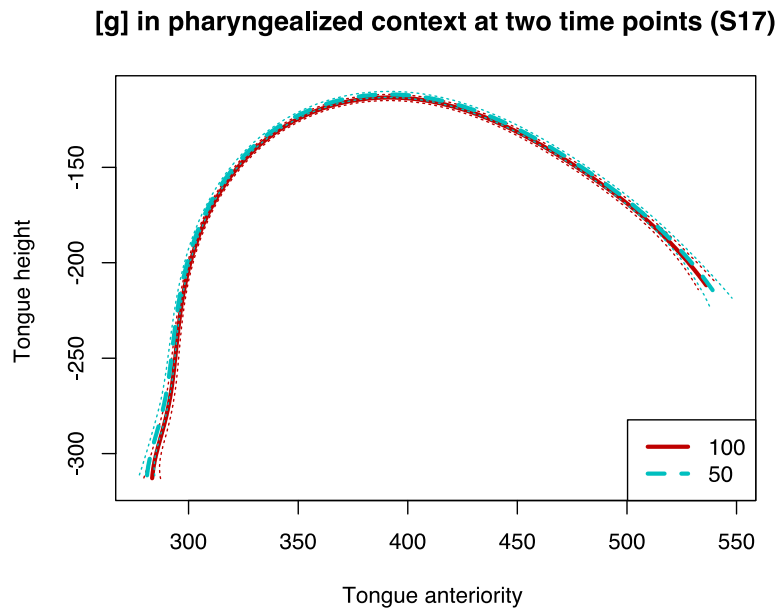


Figure 5.23. SS-ANOVA fit of [g] in pharyngealized context at two selected time points.

5.4.3.2 Resistance and compromise in [χ]. There was marginal dorsum lowering in [χ] when occurring before C^ɣ (Figure 5.24) but this lowering was variable across speakers with some speakers showing no change in dorsum posture in pharyngealized contexts. It seems that this reduction in dorsal height was operated to allow coarticulatory root retraction to take place, which partially supports the compromise hypothesis postulated earlier in this chapter. No change in dorsal height or root position was observed among other speakers such as the one presented in Figure 5.25; however, in some resistance cases, the root in producing [χ] in both contexts assumed the same retracted position as in C^ɣ.

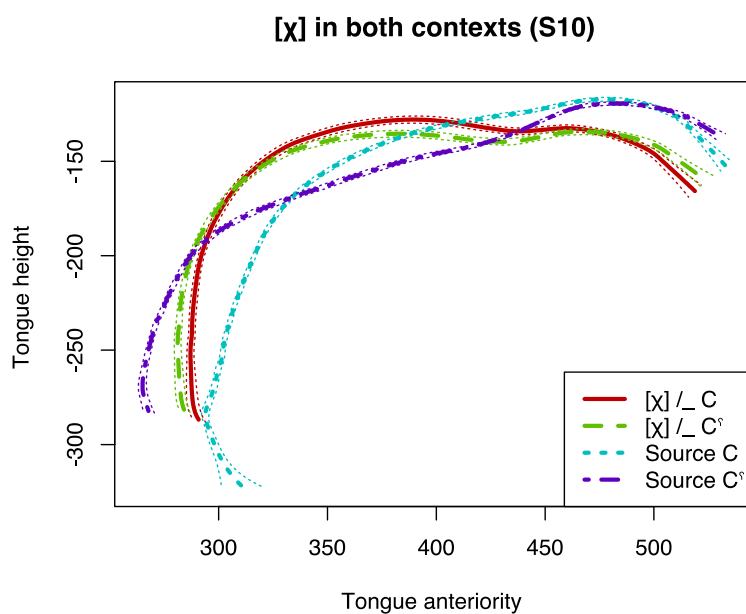


Figure 5.24. SS-ANOVA fit at mid-closure and 95% CI showing velar [χ] slight dorsal lowering in pharyngealized context and distant root position from C^ϵ .

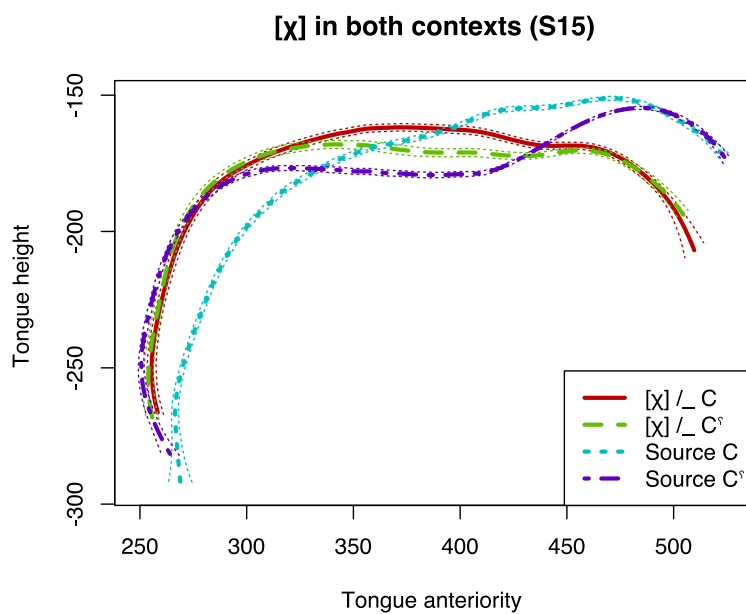


Figure 5.25. SS-ANOVA fit at mid-closure and 95% CI showing resistance of velar [χ] and its proximity to the root position of C^ϵ .

As illustrated in the scatterplot in Figure 5.21 (left), compared to [χ] in plain context, coarticulatory effect on [χ] yielded an extremely small backward displacement of the tongue root across many speakers and a small lowering of the tongue body. The tongue root retraction and body lowering were larger among a few other speakers. As in Figure 5.21 (right), relative to the source C^s, [χ] in pharyngealized context differed from the source C^s in the position of the tongue body to a greater extent than at the tongue root. The uvular [χ] was produced with a high tongue body position compared to its position in articulating the source C^s. [χ] and the source C^s were produced with a similar tongue root retraction by some speakers, but with more retraction in [χ] by other speakers. The raising of the tongue body and dorsum and the retraction of the tongue root in the articulation of [χ] provides support that this consonant is uvular rather than velar.

The crucial question here is whether [χ] and [g] reacted differently to pharyngealization given that they involve different degrees of gestural conflict, and this question can be addressed by comparing the distribution of speakers' data points in Figure 5.21. Relative to their variants in plain context, both [χ] and [g] exhibited the same range of tongue root backward displacement when occurring in a pharyngealized context, and this displacement was very small in the two consonants. However, the two velars differed in how distant their articulation was from the source C^s. That is, the root was substantially more advanced and the body was higher in [g] than C^s whereas the tongue root in [χ] was slightly more retracted and the body was higher than in C^s. Basically, [g] had an advanced tongue root but [χ] had a retracted tongue root, but this retraction was inherent in the production of [χ] regardless of the context.

The root retraction ratio of [χ] in a pharyngealized context was slightly higher than the ratio of the same consonant in a plain context. This small retraction was mostly associated with dorsal lowering and was true for some speakers only – the one in Figure 5.26 is an example. The small

difference between the retraction ratios of the two $[\chi]$'s by the speaker shown in Figure 5.24 and 5.26 was reflected in its small F-value (ANOVA: $F(1,38) = 8.08$, $p > 0.001$). There were about 4 speakers demonstrating this extent of retraction in a pharyngealized context. The rest of the speakers, like the one in Figure 5.27, exhibited an even smaller difference between the two $[\chi]$ conditions in terms of their retraction ratios. This close proximity, which signals the resistance of $[\chi]$ to coarticulatory root retraction, was consistent among all speakers except one (S22 who produced $[\chi]/_C^s$ with gradient coarticulatory root retraction compared to $[\chi]/_C$). For example, speaker S15 whose SS-ANOVA and retraction ratios are shown in Figure 5.25 and 5.27 exhibited an extremely marginal difference between the $[\chi]$ conditions: $F(1,36) = 0.56$, $p > 0.001$.

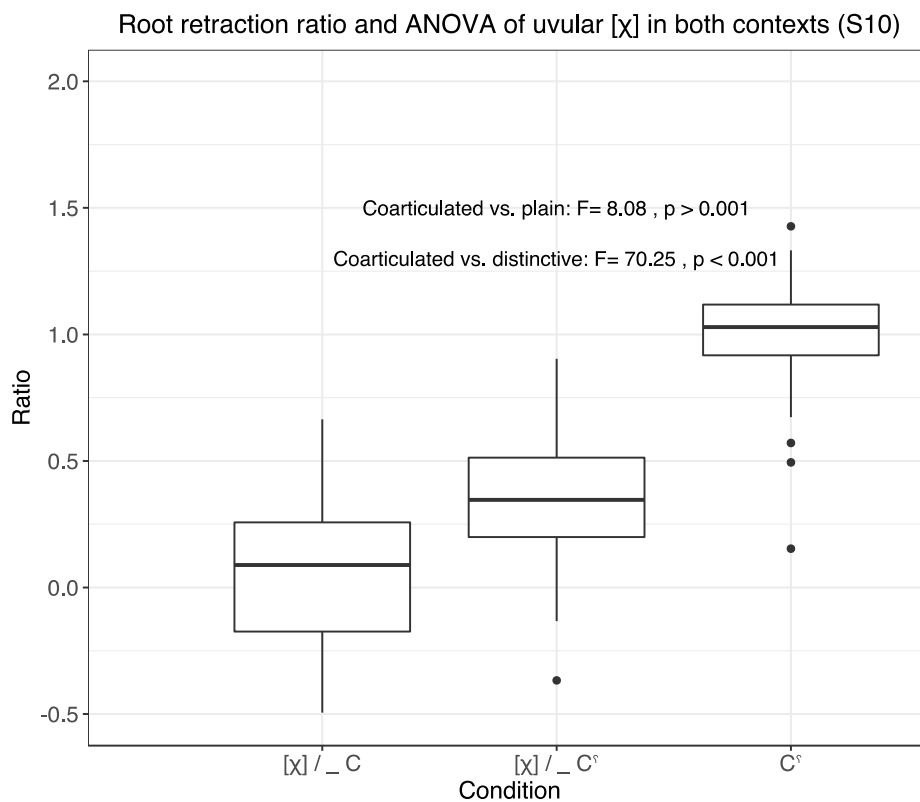


Figure 5.26. ANOVA and ratio of mean root radius showing uvular $[\chi]$ in both contexts showing higher retraction ratio before C^s .

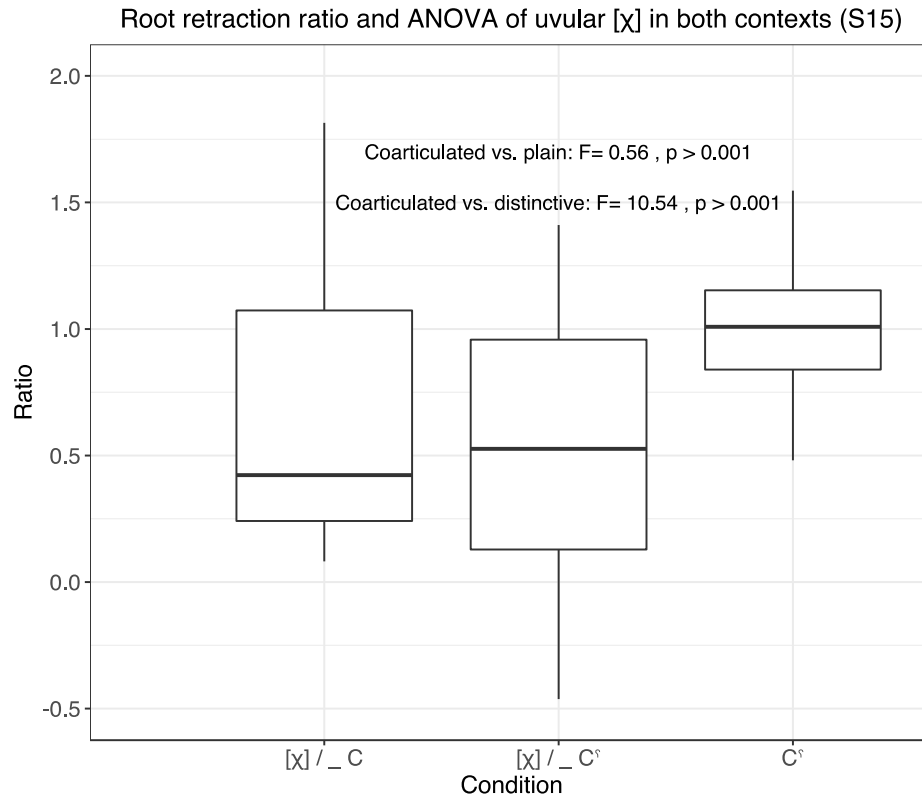


Figure 5.27. ANOVA and ratio of mean root radius showing uvular [χ] in both contexts showing similar retraction ratio before C^s and C.

5.4.4 Comparison of all target consonants. The initial categorization of the target consonants when designing the study was based on the place of articulation: labials as a control, palatals representing extreme antagonism with their tongue root advancement (same-articulator conflict), and velars representing moderate inter-articulator conflict between the dorsum and the root. However, results in this study suggested a different categorization of these consonants based on the tongue shape of the canonical phone. That is, the two consonants [χ g] had different tongue configurations and, consequently, imposed different constraints on coarticulatory pharyngealization.

Inspection of Figure 5.28 which is representative of the entire study sample (except for the root retraction in [ʃ] which was greater among other speakers) reveals three categories of the target consonants based on the canonical tongue shape in non-pharyngealized context. First, [g ɟ j] all were characterized by an advanced tongue root accompanied by tongue dorsum and body raising. They all share a similar tongue shape although they varied in the degree of palatal constriction with [g] having the narrowest and most posterior constriction and [j] having the highest tongue body and most anterior constriction. The set of target consonants which requires two antagonistic gestures, tongue root advancement and dorsum raising, [g], [ɟ] and [j], were labelled Advanced-Raised and this included [g], [ɟ] and [j]. In contrast, [χ], which turned out to be uvular, has a unique tongue shape with a relatively high and retracted tongue dorsum forming a constriction more posterior than the Advanced-Raised consonants. It forms its own category, which I labelled the Raised-only, referring to the dorsum position as the only gesture conflicting with pharyngealization. Finally, the labials [b f] in which the tongue root is in a neutral position, neither advanced nor retracted, compose a third category I call Neutral. The height of the tongue body in labials is also neutral, neither raised as in [g] nor lowered as in C^ʕ.

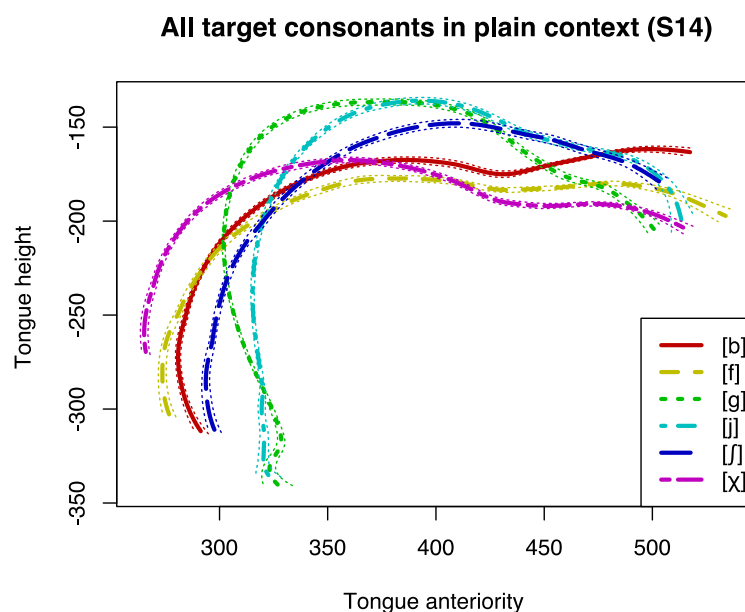


Figure 5.28. SS-ANOVA fit at mid-closure and 95% CI showing labials, palatals and velar/uvular in plain context.

Figure 5.29 shows the tongue fit of all target consonants in the pharyngealized context. The most notable finding was twofold: the root position of the Advanced-Raised consonants [g ʃ j] and the Raised-only [χ] was unchanged in the pharyngealized context, but it was retracted backward in articulating the unconstrained labials [b f]. The extent of coarticulatory retraction in [b f] was either less than (or the same) as in [χ]/_C.

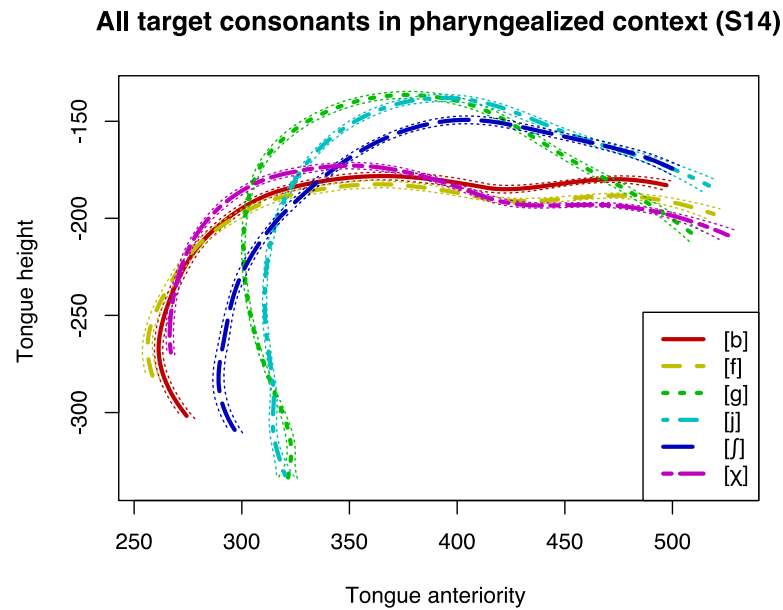


Figure 5.29. SS-ANOVA fit at mid-closure and 95% CI showing labials, palatals and velar/uvular in pharyngealized context.

5.5 Discussion

5.5.1 Gestural antagonism and phonemic pharyngealization. In phonemic pharyngealization, the pharynx is constricted by drawing the anterior pharyngeal wall, the tongue root, backward towards the back pharyngeal wall. This backward retraction, by virtue of the tongue's volume preservation property, causes a flattening and lowering of the anterior part of tongue body along the mid-sagittal plane.

Most importantly, phonemic pharyngealization is not affected by the preceding consonant even if the consonant is highly antagonistic. No carry-over coarticulation is found in /tʰ sʰ/ causing a reduction in their root retraction or tongue body depression when preceded by consonants with an advanced root such as [g ʃ j]. This contradicts the DAC model which predicts that highly constrained consonants, such as palatals in this case, should not only resist coarticulatory effects

from neighboring segments but also impose their own coarticulatory influence. Based on this prediction, the effect of the preceding palatal should be realized in producing the source C^s with a tongue body higher than the canonical form and, thus, with less retracted tongue root. This was not found in the present study, suggesting no coarticulatory dominance from the part of palatal and velar consonants.

5.5.2 Coarticulatory tongue root retraction. Coarticulatory pharyngealization on unconstrained consonants preceding the pharyngealization source but separated by a low vowel [a a:], induces a tongue configuration similar to some extent to that of the source C^s. Specifically, the bilabial /b/ and the labiodental /f/, which are not specified for any tongue root movement (i.e. unconstrained), are articulated with a more retracted tongue root and a flatter and lower tongue body than their canonical non-pharyngealized variants. The process is phonetic rather than phonological given that the magnitude of their coarticulatory gestures (e.g. /f/ in /fa:tʰna:t/ ‘smart’) is less extreme than that of phonemic pharyngealization (e.g. /tʰ/ in /fa:tʰna:t/ ‘smart’) which has a more constricted pharynx, lower and flatter tongue body. A few speakers present an exception to this conclusion as their assimilatory root retraction is as extreme as in articulating phonemic pharyngealization, which suggests that the effect of pharyngealization can be phonological in the absence of gestural constraints. Also, the two labials examined here show the same degree of coarticulatory pharyngealization and this effect extends as far as the 25% point of /f/ and mid closure of /b/.

The study examined two degrees of articulatory conflict with pharyngealization: a same-articulator and an inter-articulator conflict. The former, which refers to the tongue root advancement as in [g ʃ j] and is in direct antagonism with coarticulatory pharyngealization, poses extreme articulatory constraints on the consonants’ sensitivity to coarticulation. This contributes

to their resistance to coarticulation, whose purpose is to maintain their canonical gestures at the tongue root and dorsum. In fact, retracting the tongue root in this case would require stretching the tongue beyond its physical capacity (thus, violating its volume preservation characteristic) unless the dorsum is lowered or pushed backward to make a more posterior palatal constriction and allow coarticulatory root retraction to take place. This yields support to the resistance hypothesis and is at odds with the compromise hypothesis which predicted a reduction in the antagonistic dorsal gesture, but the dorsal constriction should still fall within the articulatory space of the consonant in order not to lose its identity. For example, similar to the case of posterior versus anterior velar constriction in [ki] vs. [ka], I expected a more posterior palatal constriction in [g ʃ j] in addition to a reduction in tongue dorsum/body raising when occurring in a pharyngealized context. However, the finding on these consonants' response to articulatory conflict does not confirm the compromise hypothesis as the dorsal articulation was sustained completely in the pharyngealized context.

The fricative [χ] presents another case of gestural antagonism. It should be noted here that this consonant is traditionally classified as velar (and transcribed as [x]), as illustrated in Table 1.3; however, ultrasound data presented here have shown otherwise: it is rather a uvular consonant as it is articulated with dorsal raising and retraction accompanied with tongue root retraction. This retraction is inherent in the consonant and is not affected by the upcoming pharyngealized consonant: [χ]/_C and [χ]/_C^s are both produced with the same tongue root displacement. This moderate antagonism between the dorsal raising of [χ] and coarticulatory pharyngealization results in a slight reduction in the dorsal gesture of [χ] when preceding a pharyngealized consonant compared to the same consonant in a non-pharyngealized context. This slight lowering of the dorsum, which partially supports the compromise hypothesis, seems to allow some coarticulatory retraction at the tongue root which it is variable across speakers.

CHAPTER 6

General Discussion

6.1 Introduction

This dissertation addresses three aspects of emphasis in Arabic: first, the articulatory realization of emphasis as pharyngealization or uvularization; second, the extent to which its assimilatory effect on the preceding consonants is susceptible to the influence of phonological and physiological constraints; and third, whether this effect is phonetic or phonological. This investigation was primarily motivated by a need to complement the existing literature, which only investigated the effect of emphasis spread on vowels. It also seeks to further our understanding of whether the counteraction between the constraints and coarticulation that have been documented in other languages can be found in emphasis spread in Arabic. To achieve these goals, ultrasound data were obtained from native speakers of Eastern Peninsular Arabic while articulating a set of consonants occurring in an emphatic and a plain context. This allows an examination of the tongue shape along the mid-sagittal plane and a comparison of the magnitude of the displacement of different tongue regions (mainly the tongue body, dorsum and root) in different phonetic contexts (such as in the presence and absence of an emphatic trigger).

As already discussed in Chapters 4 and 5, the phonological and articulatory models chosen to interpret the results obtained in this dissertation are the target-interpolation model ([Cohn, 1990](#))

and the degree of articulatory constraints model ([Recasens et al., 1997](#)), respectively. Therefore, the results of this dissertation are discussed in relation to these models with a special emphasis on the (in)consistencies between the models' predictions and the results.

6.2 Lingual Configuration in Producing Emphatics

The articulatory realization of emphasis is controversial among linguists as some laboratory evidence has suggested that the principal articulator involved in the posterior constriction is the dorsum – uvularization ([Al-Ani, 1970](#); [Bin-Muqbil, 2006](#); [Ghazeli, 1977](#); [Zawaydeh, 1997, 1998, 1999](#); [Zawaydeh & de Jong, 2011](#)) – while other evidence has shown that the constriction is formed at the tongue root – pharyngealization ([Catford, 1977](#); [Embarki, Ouni, Yeou, et al., 2011](#); [Hassan & Esling, 2007, 2011](#); [Zeroual et al., 2011](#)).

Articulatory data presented in this dissertation have revealed that emphasis is a form of pharyngealization rather than uvularization. This lingual configuration was observed in the articulation of both the emphatic and the assimilated consonants (yet with a different extent of tongue displacement). The study therefore advocates the use of the terms 'pharyngealization' and 'pharyngealized consonant' instead of 'emphasis' and 'emphatics', respectively, since the former terms reflect more accurately the articulatory mechanism responsible for producing these consonants. The evidence in favor of pharyngealization is based on two articulatory parameters: specifically, the location of the maximal posterior lingual displacement and the overall tongue shape. They demonstrate that it is the tongue root that undergoes the most salient movement in forming the secondary pharyngeal constriction. The backward retraction of the tongue root towards the posterior pharyngeal wall is a predictor of a cross-sectional compression of the pharyngeal tube (although the ultrasound imaging here shows only the tongue root movement and not the size of the pharynx). Other lingual movements observed in this study seem to be consequences of the

movement of the tongue root caused by the hydrostatic nature of the tongue and the anatomical coupling between the root and the adjacent tongue regions. They include the depression and flattening of the tongue body and a slight backward retraction and lowering of the dorsum to assist with the movement of the tongue root along the same direction. This configuration is at odds with the studies supporting uvularization which reported a combined retraction and raising of the dorsum to form a constriction against the upper pharyngeal wall. Put simply, the lowering of the tongue dorsum observed in this study leaves no evidence of any dorso-pharyngeal constriction as claimed in studies supporting uvularization. The lingual configuration during the production of phonemically pharyngealized consonants reported here forms the basis for assessing the nature of assimilatory pharyngealization discussed in the remainder of this chapter.

The fact that there is no consensus in the literature about how emphasis is realized articulatorily has led to the conclusion that the articulation of emphasis as uvularization or pharyngealization is a sort of linguistic variation. Based on a survey of the existing literature in §1.3.1.2.3, I concluded that this variation cannot be attributed to dialectal differences or reduced to methodological issues. This means that there is no systematic divide between dialects in terms of how emphasis is articulated and that it can be realized in both ways across speakers of a given dialect. The Eastern Peninsular variety examined in this dissertation seems much less variable, but it is not necessarily an exception: although the ultrasound evidence demonstrates that emphatics are pharyngealized among all speakers who participated in the study, this does not exclude the possibility that other speakers of the same dialect can produce emphatics as uvularized consonants.

Despite the lack of direct evidence, the tongue root retraction is predicted to be associated with other non-lingual gestures including, for instance, the retraction of the epiglottis (since it is attached to the root and can be pushed backward by retracting the root) and the elevation of the

larynx given the dependence of the root and larynx according to ([Esling, 2005](#)). These possible movements, in addition to the inward movement of the lateral and posterior pharyngeal walls, would contribute to increasing the compression of the pharynx cross-sectionally.

6.3 Constraints on Assimilatory Pharyngealization

This dissertation has provided novel articulatory evidence about regressive assimilatory pharyngealization on consonants preceding a phonemically pharyngealized trigger. In line with the description of emphatics in the previous section, this assimilatory effect is realized by more backward displacement of the tongue root and a lower tongue body and dorsum than in the canonical form of the respective consonants. Based on the eyetracking findings of [Alwabari and Zamuner \(2016\)](#), these assimilatory effects which could be realized acoustically are expected to be perceptible and contribute to spoken word recognition by activating lexical candidates with a phonemically pharyngealized consonant even before this consonant (C^s) is encountered.

The factors that constrain this assimilatory effect, however, are a bit more complex than what was anticipated by the initial hypotheses. Coarticulatory pharyngealization is subject to the influence of physiological constraints in ways that roughly match preliminary expectations. However, phonological constraints have unexpected effects.

6.3.1 Phonological constraints on assimilatory pharyngealization. Previous cross-linguistic studies reflect a counteraction between a language-particular system of phonemic contrast and coarticulation ([Bradlow, 1995](#); [Magen, 1984](#); [Manuel, 1987, 1990, 1999](#); [Manuel & Krakow, 1984](#); [Öhman, 1966](#)). It is assumed that coarticulation that would lead to the neutralization of phonemic contrast, and arguably allophonic contrast according to [Mok and Hawkins \(2004\)](#), is avoided. Therefore, the spatial and temporal extent of coarticulation is

restricted in order to avoid blurring the distinction between coarticulated segments and phonemes that underlyingly carry the assimilated property.

One of the goals of this dissertation was to test this claim of a counteraction against assimilatory pharyngealization in Arabic given that pharyngealized coronals /tʕ sʕ/ contrast with plain /t s/. According to Manuel's contrast hypothesis, if plain coronals /t s/ are coarticulated with the pharyngealized ones, the extent of coarticulation should be moderate, so that they remain distinct from /tʕ sʕ/. This prediction was examined by comparing assimilatory pharyngealization on plain /t s/ and another set of coronals that do not contrast in pharyngealization, /n l/, along two dimensions: the posteriority of the secondary assimilated constriction and the magnitude of articulator displacement.

The present results have shown evidence of a categorical assimilatory pharyngealization on the plain coronals /t s/, suggesting that contrast does not limit assimilatory pharyngealization. Contrary to Manuel's hypothesis, assimilation resulted in a neutralization of the contrast between /s/ and /sʕ/ in the majority of speakers (60% of them) and between /t/ and /tʕ/ in some speakers (27%). The fact that at least one consonantal contrast is neutralized is sufficient to rule out the contrast hypothesis in Arabic pharyngealization. Although /t/ is resistant to assimilation in about half the speakers, this resistance does not seem to be contingent on contrast preservation because the same pattern is observed in the coronal /n/, which does not contrast in pharyngealization. In other words, although there is no pharyngealized phoneme */nʕ/ that could be confused with an assimilated /n/ and no gestural conflict that could constrain assimilatory pharyngealization on the nasal, it is resistant to assimilation in the majority of the speakers, just like the contrastive /t/. This suggests that the resistance of /t/ and /n/ to assimilation in some speakers may be due to something other than phonological contrast and gestural conflict. This categorical pattern is not exclusive to

the consonants with phonemic contrast, /t s/. Rather, other consonants exhibit a similar pattern and, as shown in the following section, physiological constraints contribute substantially to this categoricalness. The categorical and gradient patterns which signal the presence of phonological and phonetic effects, respectively, are elaborated on in §6.4.

6.3.2 Physiological constraints on assimilatory pharyngealization. Gestural conflict, which stems from the physiological constraints on the tongue such as muscular antagonism and tongue volume preservation, can interfere with coarticulation. It can arise when two segments involved in coarticulation impose antagonistic gestures at the same articulator (e.g. [na] where the tongue dorsum raising in [n] is followed by a lowering in [a]) or at different articulators that are anatomically coupled (e.g. [na] where the tongue blade raising in [n] is followed by the lowering of an anatomically dependent tongue body lowering in [a]). There is a lack of automaticity in resolving the conflict between the coarticulation trigger and target (conflict avoidance by means of resisting coarticulation is not always the sole solution); rather, it can stimulate multiple strategies to resolve it, mainly elision, transition and compromise.

In order to examine possible strategies employed in resolving gestural antagonism with tongue root retraction, this dissertation probed different types of gestural conflict with pharyngealization. Given the articulatory evidence provided in this dissertation that the principal gesture in pharyngealization is the retraction of the tongue root accompanied by a depression of the tongue body and dorsum, three sets of consonants with different types of gestural antagonism were used in this study. Namely, the study examined coarticulation on [j ɟ g], where tongue root advancement directly counteracts tongue root retraction (same-articulator conflict); uvular [χ], where tongue dorsum raising is in an inter-articulator conflict with root retraction; and labials [b f], where there are no lingual constraints (no conflict). These consonants are grouped according to

the ultrasound findings on their primary gesture in their articulation and this grouping of consonants is different from what was expected when designing the study.

The data in this dissertation indicate that 1) the magnitude of coarticulation and the conflict resolution strategy are contingent on the degree of antagonism between coarticulatory tongue root retraction and the gestures of the target consonant; and that 2) there is a lack of a one-to-one correspondence between gestural conflict and a given reconciliation strategy. Rather, multiple strategies can be used to solve the conflict between the two segments.

The most extreme gestural conflict results from tongue root advancement (regardless of whether it is the primary or secondary gesture of the target consonant), and the least severe conflict is caused by dorsal raising. These two degrees of gestural antagonism evoke different strategies to resolve the conflict with coarticulatory tongue root retraction. The tongue root advancement in the palatals [j ʃ] and the velar [g] imposes extreme constraints on the coarticulatory tongue root retraction of pharyngealization. Consequently, such gestural antagonism leads to a complete resistance to coarticulation (i.e. a categorical effect) as the position of the tongue root and dorsum in [j ʃ g] in a pharyngealized context is not different from their position in a non-pharyngealized context. It is worth noting here that although these three consonants show a small variation in the tongue height and posteriority of their dorsal raising, this does not seem to influence their resistance to pharyngealization as all three consonants exhibit similar degrees of resistance to coarticulation.

On the other hand, the uvular [χ], which is articulated by moving the dorsum upward and backward towards the uvula while retracting the tongue root, presents another case of gestural antagonism with pharyngealization. This dorsal raising is in conflict with the secondary dorsal depression in pharyngealization, but the tongue root retraction is not antagonistic to

pharyngealization. As a result of this moderate antagonism, resistance is not the only strategy used to resolve this antagonism (unlike the antagonism in highly constrained consonants [j ʃ g]); rather, a compromise solution is also employed, in which the antagonistic dorsal raising of [χ] is reduced to allow a certain degree of coarticulatory tongue root retraction. More specifically, a canonical [χ] in non-pharyngealized context has a higher tongue dorsum position and often a less retracted tongue root than its counterpart in a pharyngealized context although the dorsal lowering is quite small. The decrease in dorsal raising can be attributed to the volume preservation characteristic of the tongue because retracting the tongue at the root area seems to pull the rest of the tongue mass downward – analogous to extending a hydrostatic body, e.g. a balloon filled with water, on one dimension which causes a compensatory reduction on another dimension.

The absence of gestural antagonism allows consonants to coarticulate more freely as seen in the labials [b f]. Since they do not involve any antagonistic dorsal raising or root advancement, these lingually unconstrained consonants exhibit a remarkable coarticulatory tongue root retraction and tongue body lowering compared to their canonical tongue shape in non-pharyngealized contexts. The magnitude of their tongue dorsum and root displacement is either the same or smaller than that required to produce phonemically pharyngealized consonants. In other words, the effect on unconstrained labials is either categorical (20% of speakers) or gradient (53% of them), an unexpected finding that leaves open the question of why these consonants often undergo full assimilation if partial (gradient) coarticulation would be sufficient.

The fact that gestural antagonism in coarticulation of pharyngealization can be solved by either resistance or reduction raises the question of why the conflict with coarticulatory tongue root retraction does not induce other reconciliation strategies observed cross-linguistically. The case of gestural antagonism between retracted /q sʰ/ and the high tense vowel /i/ in Chilcotin and

Nuu-chah-nulth reported in an earlier chapter ([Gick & Wilson, 2006](#)) is analogous to the antagonism between retracted tongue root in pharyngealization and the high consonants [g ʃ j] tested in this study. However, the reconciliation strategies used in the two cases are completely different, and this difference is not arbitrary but can be attributed to two reasons. First, the antagonism in Chilcotin and Nuu-chah-nulth involves a vowel as the coarticulation target whereas in the present study the target is a consonant with a tighter oral constriction than /i/. According to the DAC model, vowels that normally do not require an oral constriction are less constrained and, therefore, less resistant to coarticulation than the consonants used in this study. This means that the vowel can be reduced (i.e. lax) in response to coarticulation but the tongue root advancement in [g ʃ j] cannot. Second, the excrescent schwa reported in Chilcotin and Nuu-chah-nulth appears between two consecutive, conflicting consonants. In the present study, however, the conflicting consonants are not immediately adjacent but separated by the inter-consonantal vowel [a a:]. Although this vowel is not epenthetic, it might serve a similar transitional role as the excrescent schwa in the other two languages. Beyond the type of structure investigated in this dissertation, vowel insertion seems to be a strategy that could be used in Eastern Peninsular Arabic to resolve the gestural incompatibility between immediately adjacent consonants such as /jtʕ/. If it existed, the tongue root would have to move abruptly from one extreme position to another extreme one, assuming a position similar to that of a schwa in its transition from the advanced to the retracted position. However, as far as I know, this sequence is not attested in Eastern Peninsular Arabic, which means that vowel insertion may not be used as a reconciliation strategy.

An attested strategy for resolving articulatory conflict in adjacent segments is reported in [Alwabari \(2018a, 2018b\)](#): laxing of an antagonistic vowel. These studies examined the blocking induced by palatals /i j ʃ/ on coarticulatory pharyngealization in Arabic words containing the

sequence /aCiC^s/, /aCjC^s/ and /aCjC^s/. Acoustic evidence has shown that /a/ undergoes more coarticulatory pharyngealization when separated from the pharyngealization trigger by /i/ than when separated by /j/. This suggests that the vowel /i/ allows more coarticulation than the palatal consonants used in this dissertation although all are linguistically constrained and antagonistic to pharyngealization. This finding, combined with the acoustic evidence of [Jongman et al. \(2011\)](#) that /i/ can be pharyngealized (although to a lesser extent than other vowels), suggests that /i/ in Arabic is lax (lowered) as in Chilcotin and Nuuchahnulth. In brief, although /i/ has a similar tongue shape and gestural antagonism as [g ʃ j], it seems to induce a different reconciliation strategy than [g ʃ j]. The fact that [g ʃ j] respond to antagonism by resisting coarticulation and not by reducing their tongue body height as in [x] and [i] (in Arabic and the two other languages) could be attributed to the severity of the articulatory constraints that affect them.

6.4 Phonetics and Phonology of Assimilatory Pharyngealization

A central issue of this dissertation is the classification of the contextual effect exerted by the pharyngealized consonants as either phonetic coarticulation or phonological assimilation. In fact, the study was designed to address this issue by examining assimilation on /t s/ using the experiment reported in Chapter 3; however, the other data of this dissertation obtained to address the other two issues (i.e. the phonological and physiological constraints reported in Chapter 4 and 5, respectively) make a significant contribution to answering the question of whether the effect is phonetic or phonological. Given this, the discussion here probes the findings of all consonants tested in this dissertation.

In an earlier chapter, I established certain criteria for assessing whether the contextual effect of pharyngealization is phonetic or phonological. Namely, the conceptualization of the effect as either phonetic or phonological is conditioned by its gradient or categorical nature both

qualitatively (in terms of the overall tongue configuration and the posteriority of the assimilatory gesture relative to phonemically pharyngealized consonants) and quantitatively (i.e. the magnitude of tongue displacement).

The results suggest that the phonetic or phonological nature of contextual pharyngealization is largely dependent on the presence of phonological and/or physiological constraints and on the severity of the latter. Due to these constraints, the target consonants in this dissertation range in their susceptibility to pharyngealization from categorical assimilation to complete resistance, with some intermediate degrees of gradient coarticulation. Table 6.1 illustrates the type of contextual effect on all target consonants along with the constraints that characterize them and shows the chapter in which the results of the respective consonant are reported.

Table 6.1

Summary of constraints on all consonants and the nature of contextual pharyngealization in each

C	Constraints		Effect	Chapter
	Phonological	Physiological		
[t]	Yes (contrast)	No	Resistance or assimilation	3-4
[s]	Yes (contrast)	No	Assimilation	3-4
[n]	No	No	Resistance	4
[l]	No	No	Coarticulation or assimilation	4
[b]	No	No	Coarticulation or assimilation	5
[f]	No	No	Coarticulation or assimilation	5

[χ]	No	Yes (high dorsum)	Resistance (and reduction)	5
[g]	No	Yes (advanced root)	Resistance	5
[ʃ]	No	Yes (advanced root)	Resistance	5
[j]	No	Yes (advanced root)	Resistance	5

Three major conclusions can be drawn based on this survey of consonants presented in the table above. Firstly, the presence of either type of constraints, phonetic or phonological, enhances categoricalness. Avoiding gradience and permitting only categoricalness in certain consonants (which either undergo substantial assimilation or resist it), owes to the presence of either a phonemic contrast as in /t s/ or physiological constraints as in [χ g ʃ j].

However, while physiological constraints force resistance exclusively, phonological ones lead to both extremes of categoricalness: resistance and assimilation. It is not surprising that gesturally constrained consonants only resist coarticulation and do not undergo any coarticulatory or assimilatory effect, as this is consistent with several cases reported in the literature (discussed in details in §2.3.2). The effect of phonological constraints does not seem to be arbitrary. In order to preserve the contrastive structure in the language (i.e. the contrast between plain /t s/ and pharyngealized /t^ɕ s^ɕ/), assimilation leads to a phone that must have the same phonetic characteristics of the contrastive phonemes present in the language. This means that maintaining distinctiveness is not achieved exclusively by resisting coarticulation as Manual claims, but by restricting the assimilatory effect in such a way as to produce a phone with the same canonical form of distinctive phonemes: a phone that is similar to either plain /t s/ by resisting coarticulation or pharyngealized /t^ɕ s^ɕ/ by undergoing assimilation. For this reason, partial coarticulation in /t s/ is not allowed.

The last conclusion from this survey is that the absence of constraints, as in labials and lateral /l/, does not automatically lead to gradience but allows a more flexible reaction to pharyngealization. The absence of both phonemic contrast in pharyngealization and gestural conflict in the labials [b f] and coronals [n l] leads to gradient coarticulatory pharyngealization in many cases. It is gradient because the magnitude of the coarticulated gesture is less extreme than that of the phonemically pharyngealized trigger. The claim made here does not mean that the absence of constraints forces the effect to be exclusively gradient, but rather it allows the consonant to freely undergo gradient coarticulation besides the categorical effects.

The phonological processes found in this study whereby /s t/ assimilate to a pharyngealized consonant is similar to a phonological process in the language whereby /t s/ become /tˤ sˤ/ in certain morphological contexts containing a pharyngealized consonant. This process, known as [ʔɪbda:l] ‘consonant mutation’, occurs in words with a pharyngealized consonant (the initial consonant of the word root) preceding a /t/ that is part of the morphological template. In this case, the templatic /t/ is changed into /tˤ/ ([see Merrill, 2014 for a survey of mutation in several languages](#)). For example, the pharyngealized [tˤ] in [əsˤtˤaħab] ‘he/it chose’ and [əsˤtˤabar] ‘he/it is patient’ is underlyingly /t/ but altered into [tˤ] due to the preceding phonemically pharyngealized /sˤ/. This alternation is due to progressive, local assimilation. The underlying forms of these two examples are /əsˤtaħab/ and /əsˤtabar/, respectively, with plain /t/ in both. This consonant alternation is triggered in certain morphological conditions and becomes phonologized as the underlying /t/ is lost. The alternation applies to all trilateral roots starting with an initial pharyngealized consonant /tˤ sˤ dˤ ðˤ/ (e.g. the roots ‘sˤ-ħ-b’ and ‘sˤ-b-r’ of the previous example words) in specific morphological templates such as ɪCtaCaC (where the three Cs are the root consonants). Other

similar morphological templates (but not all) trigger the same alternation and some of these templates are used in Eastern Peninsular Arabic.

6.5 Pharyngealization and Models of Coarticulation

The results of this dissertation have significant implications for models of coarticulation – namely, the target-interpolation model proposed by [Cohn \(1990\)](#) and the DAC model developed by [Recasens et al. \(1997\)](#). As explained in an earlier chapter, the two models do not only differ in their basic input unit, but also in whether they account for phonological or phonetic processes. Similar to §6.4, the discussion of the two models will not be limited to a subset of the data obtained in this dissertation (e.g. limiting the DAC model to the findings of Chapter 5); rather, the results of all consonants are collectively discussed in relation to the two models. The inclusion of all results raises concerns about the models that are worth mentioning. The discussion here will focus on two main elements: areas of (in)consistencies between the present findings and the models, and the aspects of the models that are beyond the scope of this dissertation.

6.5.1 Pharyngealization and the target-interpolation model. The target-interpolation model predicts that segments unspecified underlyingly can be subject to two levels of segmental processes: phonological and phonetic, with the former taking precedence over the latter. An underlyingly unspecified segment can categorically assume the feature value of the trigger segment by means of a phonological rule such as feature spreading (if there is any feature spreading in the language). The target segment would thus become specified in the output of phonology. Phonetic implementation takes place only if the segment is still unspecified phonologically (e.g. if no feature spreading has taken place), and the output of phonetics can be either cline-like (gradient) or plateau-like (categorical) within the limits of phonetic constraints.

6.5.1.1 Results raising concerns about the model. The model allows us to account for the varying effects on all consonants surveyed in Table 6.1, whether the effects are assimilatory through application phonological rules or coarticulatory via phonetic interpolation. Nevertheless, the fact that some consonants undergo two types of effects poses serious problems for the model. If a phonological rule such as feature spreading exists in the language, then why is this rule applied to a given consonant by some speakers and not by other speakers, and why do they apply it to some consonants but not others? The model is not constrained enough in the sense that phonological rules can be invoked ad hoc to explain any segmental effects.

To illustrate this issue, consider the pharyngealization effect on the labials [b f] and the lateral [l]. The present ultrasound findings reveal that these consonants, which are underlyingly unspecified for pharyngealization and unconstrained articulatorily, undergo categorical assimilation for some speakers and gradient coarticulation for others. This means that, according to the target-interpolation model, they are pharyngealized either at the phonological level via feature spreading or at the level of phonetic implementation via interpolation. This raises the questions of what rule governs this variability among speakers and why such rule allows the feature spreading rule in some but not all speakers. This is particularly problematic when considering that there are apparently no phonological or phonetic constraints that could intervene with either feature spreading or phonetic interpolation. The contrastive /t/ presents a quite similar case. This consonant is assimilated for some speakers and resistant for others. This means that in some cases, this consonant is subject to feature spreading by which the underlying [-RTR] is changed to [+RTR] at the output of phonology, and that in other cases the consonant remains [-RTR] at the output of phonology. This, again, raises the issue of why the feature spreading rule is not applied in /t/ consistently.

Another substantial issue with the model is that, different consonants that have the same phonological specification and lack phonetic constraints on coarticulation react differently to contextual pharyngealization. An evident example is the two coronals [n] and [l], both of which are underlyingly [ØRTR] and unconstrained phonetically; however, the nasal is resistant to pharyngealization whereas the lateral is subject to either feature spreading or phonetic implementation. The target-interpolation model can easily come up with a variable rule that accounts why [l] undergoes phonetic or phonological processing whereas [n] does not. However, this rule is merely descriptive in that it neither derives nor is constrained by principled factors.

The examples provided here illustrate other drawbacks of the target-interpolation model. First, it does not account for the lack of a one-to-one correspondence between features and phonetic parameters. The model also predicts different patterns of coarticulatory pharyngealization depending on what feature is used to represent pharyngealization: several features have been proposed in the literature to represent pharyngealization and a segment such as [k] that is unspecified for a given feature (e.g. [ØRTR]) can be specified for another pharyngealization feature (e.g. [-low]); therefore, the same segment can be subject to phonetic interpolation in the first case but not in the latter.

6.5.1.2 Evidence supporting the model. Less controversial cases of pharyngealization are the consonants with no speaker variability and whose assimilatory behavior can be readily accounted for by the model.

The resistance to assimilation in the constrained segments [χ ɡ ʃ j] provides support for Cohn's model. According to the model, phonetic constraints can pose limitations prohibiting coarticulation (via phonetic interpolation) or determining its magnitude, for example, voiceless stops in French must be oral for most of their duration which constrains nasal coarticulation. The

gestural conflict induced by advancing the tongue root and raising the tongue body in [g ʃ j] and by raising and retracting the dorsum in [χ] can be deemed as a phonetic constraint against coarticulation. Given this, the phonetic output of the interaction between this phonetic constraint and feature underspecification is a segment with no tongue root retraction. It should be pointed out, however, that although the phonetic constraints in Cohn's model are not explicitly physiological in nature (as those constraining [χ g ʃ j] are), the model is flexible enough to accommodate them.

Other supporting evidence is the categorical assimilation of [s], which is phonologically specified for the pharyngealization feature [+RTR]. The model predictions of the findings on [s] are presented in (1), and for simplicity, the schematization is presented for the CVC^s sequence only and not the entire word /sat^sarha:/ 'he hit her on the face'. For reasons explained in an earlier chapter, I choose the feature [RTR] to formally represent pharyngealization and use it for model illustration. I also use tongue root retraction as the phonetic parameter for assigning values to phonetic targets. As schematized in (1), although [s] is underlyingly [-RTR], it was observed to undergo categorical assimilation, suggestive of the role of feature spreading; therefore, this consonant leaves the phonology with [+RTR] value. With this feature specification, [s] is mapped to a phonetic target with a retracted tongue root.

1) The model's account for the effect in /sat^s/ of /sat^sarha:/

a) Phonology

a. Underlying representation

s	a	t ^s
-RTR	ØRTR	+RTR

b. Phonological rules: Feature spreading in [s] triggered by [t^s]

c. Output of phonology

s	a	tʰ
+RTR	ØRTR	+RTR

b) Phonetics

a. Target assignment:

		+RTR	+RTR	+RTR
Root retraction	Retracted	s	a	tʰ
	Neutral			

b. Phonetic constraints: No constraints

c. Interpolation:

		+RTR	+RTR	+ RTR
Root retraction	Retracted	s	a	tʰ
	Neutral			

Lastly, the model accounts for coarticulation not only in a segment that is immediately adjacent to the trigger, but also in segments that are separated from the trigger by an unspecified segment. An example of non-local nasal coarticulation over two unspecified segments that the model accounts for is the case of French *maille* /maj/ ‘stitch’ where the vowel (due to [-nasal] Deletion) and the glide are unspecified and undergo a cline-like coarticulation. This means that nasalization continues in a decreasing manner throughout the two segments /aj/ ([Cohn, 1990, pp. 175-178](#)). A similar example provided in this dissertation is the non-local coarticulatory

pharyngealization observed in sequences like [bat^ɕ]). The study found evidence that the magnitude of coarticulatory root retraction in /b/ is smaller than the phonemic retraction in /t^ɕ/). Yet, still lacking is an examination of the magnitude of pharyngealization in the vowel showing whether there is an increase in pharyngealization starting from /b/ to /a/, which would conform to the target-interpolation model's prediction.

6.5.2 Pharyngealization results and the DAC model. With only a few exceptions, the coarticulation and resistance patterns of most consonants examined in this study are accounted for by the DAC model, as explained below. There are, however, certain predictions of the model that are not supported by the findings in this dissertation; and those are highlighted in this section.

A major prediction of the DAC model supported in this study is the correlation between the degree of lingual constraints and the magnitude of coarticulation. The consonants studied here can be classified into three groups according to the degree of constraints at the tongue root and dorsum: lingually unconstrained [b f t s l], moderately constrained [χ] and highly constrained [g ʃ j]. In support of the model, the unconstrained set of consonants is substantially prone to coarticulation, yet to varying degrees. The highly constrained [g ʃ j] have a complex tongue configuration which requires two simultaneous gestures, tongue root advancement and tongue body raising, both of which are antagonistic to pharyngealization. This complexity makes [g ʃ j] highly constrained and, consequently, completely resistant to coarticulation. According to the DAC model, the degree of coarticulation sensitivity/resistance is correlated with the degree of lingual constraints. Given this, and since [g ʃ j] are more constrained than [χ] which involves dorsal raising but not root advancement, the uvular [χ] tends sometime to exhibit less resistance than [g ʃ j] and often reduction in its dorsal raising.

What violates the model predictions is the variation among consonants of the same level of lingual constraint in terms of the magnitude of their coarticulation. An evident example of this is the assimilatory patterns observed in labial [b f] versus coronal [l], where no phonological constraints may contribute to categorical assimilation. Compared to labials which do not involve any lingual activity, there is a certain degree of lingual movement in the articulation of the coronal [l]; therefore, according to the DAC model, the coronal [l] should be less susceptible to pharyngealization than the labials [b f]. This is at odds with the present findings. The lingual [l] is found to be more susceptible to coarticulation and categorically assimilated (by some speakers) than the unconstrained labials, which violates the model predictions. This gradient-vs-categorical nature of pharyngealization, in cases where no phonological factors are at play as in labials and [l], is beyond the scope of the DAC model. Moreover, in disagreement with the DAC model, where coarticulation is inversely related to the degree of constriction, the velar stop [g] does not present more acute resistance than the other two consonants with an advanced root although it has the highest tongue body. Rather, although the three constrained consonants differ slightly in terms of tongue body height and anteriority, they all demonstrate comparable degrees of resistance.

Another violation of the model is the lack of correlation between the resistance of [g ʃ j] and their prominence (so-called “aggressiveness”). Tongue root retraction in phonemic pharyngealization is unaffected by the presence of these consonants with an advanced tongue root, contrary to what the model assumes.

Finally, one aspect of the DAC model is not investigated in this study: the temporal extent of coarticulation. The model assumes that the degree of lingual constraints in a segment is correlated not only with the magnitude of its susceptibility to coarticulation but also with the temporal extent of coarticulation. For example, coarticulation in the moderately constrained [χ]

should be less temporally persistent than the unconstrained [f]. This claim has to be verified in an ultrasound experiment with higher temporal resolution than the current experiment.

6.6 Future Directions

The discussion in this chapter has pinpointed some gaps in our understanding of coarticulation resistance in general and coarticulatory pharyngealization in particular, and I highlight them here. First, given that the rationale for the counteraction of phonemic contrast against assimilation is the drive to maintain distinctiveness and produce clear unambiguous speech, there could be cases where assimilatory pharyngealization might lead to ambiguity and is thus avoided. I assume that contrast in this dissertation has led to some flexibility where the contrastive consonants were either categorically assimilated or resistant and I assume neutralization of contrast is allowed because it does not lead to any lexical ambiguity in the words examined here. There could be cases, however, where a neutralization of pharyngealization contrast can cause lexical ambiguity where a word with an assimilated consonant becomes indistinct from a similar word with a phonemically pharyngealized consonant. To illustrate this, is assimilation in [s] in a word containing the sequence sVC^s completely resisted if it makes the word ambiguous with another word in the language like s^sVC^s? Does contrast in this case, combined with the lexical factor just explained, prevent assimilation and make [s] consistently resistant to assimilation? This is a venue for future research.

In addition, since some segments are more prone to one direction of coarticulation than the other, depending on the timing of the gesture being subject to coarticulation, the present results on regressive coarticulation may not apply to progressive coarticulation. Therefore, and given the fact that there is no evidence about the timing of the radical gesture relative to the coronal one, more research is indeed necessary to uncover both the gestural timing and the resistance of progressive

(e.g. the sequence /tʰaj/ ~ /tʰab/) versus regressive (e.g. the sequence /jatʰ/ ~ /batʰ/) coarticulation of pharyngealization. The initial hypothesis here is that if the pharyngeal gesture is bound to the onset of the pharyngealized consonant, regressive coarticulatory pharyngealization should be more dominant on antagonistic gestures (i.e. less coarticulation resistance on segments, even those with moderate gestural conflict) than progressive coarticulation, and vice versa. Were the pharyngeal gesture initiated later during the consonant relative to the coronal gesture (i.e. bound to consonant offset), progressive coarticulation would be expected to more robust on constrained and unconstrained segments than regressive coarticulation.

Furthermore, examination of the inter-consonantal vowel in the present study can further verify the conclusion made above about the lack of correlation between the degree of lingual constraints in [g ʃ j] and the coarticulatory effect they exert on the following consonant – which violates the prediction of the DAC model. Examining the intervening vowel [a] acoustically and/or articulatorily and comparing it with other pharyngealized vowels can reveal if [g ʃ j] indeed hinder coarticulatory pharyngealization in the vowel. To illustrate this, the vowel in [gatʰ] undergoes two opposite effects: a retraction of the tongue root induced regressively by /tʰ/ and an advancement of the tongue root induced progressively by [g]. Comparing [a] in this word to the vowel in [batʰ] would allow us to determine if the constrained consonant [g] reduces the coarticulatory effect of pharyngealization (thus, suggesting a two-way counteraction between the constrained [g] and pharyngealized consonants, as the DAC model assumes). I collected acoustic and articulatory data of the trans-consonantal vowel in the experiments reported in Chapter 3-5, but the data were not analyzed.

6.7 Conclusion

The articulatory investigation conducted in this thesis confirms that gestural conflict stemming from the physiological constraints of the tongue is a conditional factor in determining a consonant's sensitivity to coarticulation and, moreover, in selecting an appropriate strategy to resolve such conflict. More specifically, extreme and moderate gestural conflict are associated with the two resolution strategies found in this study, conflict avoidance (i.e. resistance to coarticulation) and compromise (i.e. articulatory reduction), respectively.

Unlike physiological constraints, phonological factors such as phonemic contrast do not prevent assimilation on consonants that are contrastive in pharyngealization. Rather, despite having phonemically pharyngealized counterparts /s^ɣ t^ɣ/, assimilatory pharyngealization in /s/ and /t/ by some speakers is very extreme to the extent that it causes their contrast with /s^ɣ t^ɣ/ to be neutralized. Articulatory evidence has led to the conclusion that phonemic contrast does not prevent assimilation but rather contributes to its categorical nature where consonants are either categorically assimilated or completely resistant to assimilation (as in /t/ by some speakers).

The interaction between the two types of constraints, the phonological and physiological ones, is essential in the conceptualization of assimilatory pharyngealization as phonetic or phonological. In general, categorical effects (either categorical assimilation or resistance) are attributable to the presence of either or both types of constraints in a given consonant. For instance, consonants that are phonologically constrained such as /t s/ tend to categorically assimilate or resist pharyngealization triggered by an upcoming C^ɣ, and those that are physiologically constrained such as [j ɟ g χ] always resist pharyngealization. Phonetic gradient coarticulation, on the other hand, reflects a lack of both phonological and physiological constraints as evidenced by the

coarticulatory root retraction in [b f l] that are constrained neither phonologically nor physiologically.

References

- Al-Ani, S. (1970). *Arabic Phonology*. The Hague: Mouton.
- Al-Mozainy, H. (1981). *Vowel Alternations in Bedouin Hijazi Arabic: Abstractness and Stress*. (Doctorate of Philosophy). University of Texas at Austin, United States.
- Al-Solami, M. (2017). Ultrasound study of emphatics, uvulars, pharyngeals and laryngeals in three Arabic dialects. *Canadian Acoustics*, 45(1), 25-36. Retrieved from <https://jcaa.caa-aca.ca/index.php/jcaa/article/view/2603/pdf>
- Al-Tamimi, F., Alzoubi, F., & Tarawnah, R. (2009). A videofluoroscopic study of the emphatic consonants in Jordanian Arabic. *Folia Phoniatrica et Logopaedica*, 61(4), 247-253. doi:10.1159/000235644
- Al-Tamimi, F., & Heselwood, B. (2011). Nasoendoscopic, videofluoroscopic and Acoustic study of plain and emphatic coronals in Jordanian Arabic. In Z. M. Hassan & B. Heselwood (Eds.), *Instrumental Studies in Arabic Phonetics*. Philadelphia: John Benjamins Publishing.
- Al-Tamimi, J. (2015). *Spectral tilt as an acoustic correlate to pharyngealisation in Jordanian and Moroccan Arabic*. Paper presented at the 18th International Congress of Phonetic Sciences (ICPhS), Glasgow, UK. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/proceedings.html>
- Al-Tamimi, J. (2017). Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations. *Laboratory Phonology*, 8(1). doi:10.5334/labphon.19
- Alammar, A. (2017). *Emphasis in Zilfaawi Arabic*. (Doctor of Philosophy). State University of New York at Stony Brook, United States. Retrieved from <https://search.proquest.com/docview/1910863590?accountid=14701> ProQuest Dissertations Publishing database. (1910863590)
- Alhammad, R. (2014). *Emphasis Spread in Najdi Arabic*. (Doctorate of Philosophy). California State University, Fresno, CA.
- Ali, L. H., & Daniloff, R. G. (1972). A contrastive cinefluorographic investigation of the articulation of emphatic - non emphatic cognate consonants. *Studia Linguistica*, 26(2), 81-105. doi:10.1111/j.1467-9582.1972.tb00589.x
- Altairi, H., Brown, J., Watson, C., & Gick, B. (2017). *Tongue Retraction in Arabic: An Ultrasound Study*. Paper presented at the Proceedings of the 2016 Annual Meeting on Phonology, California, USA.
- Alwabari, S. (2018a). *Gradient Resistance to Coarticulation and Articulatory Constraints: An Ultrasound Study*. Paper presented at the 16th Conference on Laboratory Phonology, Universidade de Lisboa, Lisbon, Portugal.
- Alwabari, S. (2018b). *Palatals incompatibility with tongue root retraction: Ultrasound and acoustic evidence*. Paper presented at the Canadian Linguistic Association (CLA), Regina, Canada.
- Alwabari, S., & Zamuner, T. (2016). *Regressive Coarticulation of Pharyngealization in Arabic Spoken Word Recognition*. Paper presented at the 15th Conference on Laboratory Phonology, Cornell University, Ithaca, United States.
- Alwan, A. (1989). Perceptual cues for place of articulation for the voiced pharyngeal and uvular consonants. *The Journal of the Acoustical Society of America*, 86(549). doi:10.1121/1.398234

- Alzaaq, A. Y. (2017). *Arabic adaptation of loanwords: An empirical examination of pharyngealization and vowel epenthesis*. (MA). California State University, United States. Retrieved from <https://search.proquest.com/docview/1870037869?accountid=14701> ProQuest Dissertations Publishing database. (1870037869)
- Amerman, J. D., Daniloff, R., & Moll, K. L. (1970). Lip and jaw coarticulation for the phoneme /æ/. *Journal of Speech and Hearing Research*, 13(1), 147-161. doi:<https://doi.org/10.1044/jshr.1301.147>
- Baker, A. (2005). *Palatoglossatron 1.0*. University of Arizona, Tucson, Arizona. Retrieved from <http://dingo.sbs.arizona.edu/~apilab/pdfs/pgman.pdf>
- Barbosa, P. A., & Albano, E. C. (2004). Brazilian Portuguese. *Journal of the International Phonetic Association*, 34(2), 227-232. doi:10.1017/S0025100304001756
- Barlaz, M., Shosted, R., Fu, M., & Sutton, B. (2018). Oropharyngeal articulation of phonemic and phonetic nasalization in Brazilian Portuguese. *Journal of Phonetics*, 71, 81-97. doi:10.1016/j.wocn.2018.07.009
- Beddor, P. S., Harnsberger, J. D., & Lindemann, S. (2002). Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics*, 30(4), 591-627. doi:10.1006/jpho.2002.0177
- Bell-Berti, F., & Harris, K. S. (1977). Anticipatory coarticulation: Some implications from a study of lip rounding. *The Journal of the Acoustical Society of America*, 62(S1). doi:10.1121/1.2016047
- Bell-Berti, F., & Harris, K. S. (1981). A temporal model of speech production. *Phonetica*, 38, 9-20.
- Bell-Berti, F., & Harris, K. S. (1982). Temporal patterns of coarticulation: Lip rounding. *The Journal of the Acoustical Society of America*, 71(2), 449-454.
- Bell-Berti, F., Krakow, R. A., Gelfer, C. E., & Boyce, S. E. (1995). Anticipatory and carryover effects: Implications for models of speech production. In (pp. 77-96). New York: AIP Press.
- Bell-Berti, F., & Krakow, R. A. (1991). Anticipatory velar lowering: A coproduction account. *The Journal of the Acoustical Society of America*, 90(1), 112-123. doi:10.1121/1.401304
- Benguerel, A. P., & Cowan, H. A. (1974). Coarticulation of upper lip protrusion in French. *Phonetica*, 30(1), 41-55.
- Bin-Muqbil, M. (2006). *Phonetic and phonological aspects of Arabic emphatics and gutturals*. (Doctor of Philosophy). University of Wisconsin-Madison, ProQuest Dissertations and Theses database.
- Bladon, R. A. W., & Al-Bamerni, A. (1976). Coarticulation resistance in English /l/. *Journal of Phonetics*, 4, 137-150.
- Bladon, R. A. W., & Nolan, F. J. (1977). A video-fluorographic investigation of tip and blade alveolars in English. *Journal of Phonetics*, 5, 185-193.
- Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [Computer program] (Version 6.0.37). Retrieved from <http://www.praat.org/>
- Boudelaa, S., & Marslen-Wilson, W. D. (2010). Aralex: A lexical database for Modern Standard Arabic. Retrieved from <https://aralex.mrc-cbu.cam.ac.uk>. <https://aralex.mrc-cbu.cam.ac.uk>
- Bradlow, A. R. (1995). A comparative acoustic study of English and Spanish vowels. *The Journal of the Acoustical Society of America*, 97(3), 1916-1924. doi:10.1121/1.412064

- Braver, A. (2014). Imperceptible incomplete neutralization: Production, non-identifiability, and non-discriminability in American English flapping. *Lingua*, 152, 24-44. doi:10.1016/j.lingua.2014.09.004
- Browman, C. P., & Goldstein, L. M. (1990). Tiers in articulatory phonology with some implications for casual speech. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech* (pp. 341-376). Cambridge: Cambridge University Press.
- Bukshaisha, F. A. M. (1985). *An experimental phonetic study of some aspects of Qatari Arabic*. (Doctorate of Philosophy). The University of Edinburgh, United Kingdom. ProQuest Dissertations Publishing database. (303479607)
- Butcher, A., & Ahmed, K. (1987). Some acoustic and aerodynamic characteristics of pharyngeal consonants in Iraqi Arabic. *Phonetica*, 44, 156-172.
- Card. (1983). *A phonetic and phonological study of Arabic emphasis*. (PhD). Cornell University, United States. Retrieved from <https://search-proquest-com.proxy.bib.uottawa.ca/docview/303144025?accountid=14701> ProQuest Dissertations Publishing database. (303144025)
- Catford, J. (1977). *Fundamental Problems in Phonetics*. Edinburgh, UK: Edinburgh University Press.
- Cho, T. (2000). An articulatory study of prosodically conditioned V-to-V coarticulation in English. *The Journal of the Acoustical Society of America*, 108, 2467. doi:10.1121/1.4743096
- Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, 32(2), 141-176. doi:10.1016/s0095-4470(03)00043-3
- Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. Cambridge, London: The MIT Press.
- Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook*, 2, 225-252. Retrieved from <http://www.jstor.org/stable/4419958>
- Clements, G. N. (1987). *Toward a substantive theory of feature specification*. Paper presented at the UCLA Phonology Symposium.
- Cohn, A. C. (1990). *Phonetic and phonological rules of nasalization*. (PhD). University of California, Los Angeles. Retrieved from <https://search-proquest-com.proxy.bib.uottawa.ca/docview/303815840?accountid=14701> ProQuest Dissertations Publishing database. (303815840)
- Cohn, A. C. (1993). Nasalisation in English: phonology or phonetics. *Phonology*, 10(01), 43-81. doi:10.1017/s0952675700001731
- Cohn, A. C. (2006). Is there gradient phonology? In G. Faneslow, C. Fery, R. Vogel, & M. Schlesewsky (Eds.), *Gradience in Grammar: Generative Perspectives* (pp. 25-44). Oxford: OUP.
- Conklin, J., & Dmitrieva, O. (2017). Vowel-to-vowel coarticulation in Spanish non-words: Effects of stress and consonantal context. [- 2017/10/01]. *The Journal of the Acoustical Society of America*, 142(4), 2581. Retrieved from <https://doi.org/10.1121/1.5014439>
- Crumbie, L., Salvador, F., & Rad, A. (2019). Tongue. In Germany: Kenhub GmbH.
- Daniloff, R., & Hammarberg, R. E. (1973). On defining coarticulation. *Journal of Phonetics*, 1, 239-248.
- Daniloff, R., & Moll, K. L. (1968). Coarticulation of lip rounding. *Journal of Speech and Hearing Research*, 11(4), 707-721.

- Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *The Journal of the Acoustical Society of America*, 120(1), 407-415. doi:10.1121/1.2205133
- Davis, E. P. (1999). *Measurement and kinematic modeling of the human tongue*. (Doctorate of Philosophy). The Johns Hopkins University, United States. (304508993)
- Davis, S. (1993). Arabic pharyngealization and phonological features. In M. Eid & C. Holes (Eds.), *Perspectives on Arabic Linguistics V: Papers from the 5th Annual Symposium on Arabic Linguistics* (Vol. 101, pp. 149-162). Amsterdam/Philadelphia: John Benjamins Publishing.
- Davis, S. (1995). Emphasis spread in Arabic and grounded phonology. *Linguistic Inquiry*, 26(3), 465-498. Retrieved from <http://www.jstor.org/stable/4178907>
- de Jong, K., Beckman, M. E., & Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and Speech*, 36(2/3), 197-212. doi:10.1177/002383099303600305
- Delattre, P. (1972). Pharyngeal features in the consonants of Arabic, German, Spanish, French, and American English. *Phonetica*, 23(3), 129-155. doi:10.1159/000259336
- Delvaux, V., Demolin, D., Harmegnies, B., & Soquet, A. (2008). The aerodynamics of nasalization in French. *Journal of Phonetics*, 36(4), 578-606. doi:10.1016/j.wocn.2008.02.002
- Desmeules-Trudel, F., & Brunelle, M. (2018). Phonotactic restrictions condition the realization of vowel nasality and nasal coarticulation: Duration and airflow measurements in Québécois French and Brazilian Portuguese. *Journal of Phonetics*, 69, 43-61. doi:10.1016/j.wocn.2018.05.001
- Dow, M. (2014). *Contrast and markedness among nasal(ized) vowels: A phonetic-phonological study of French and Vimeu Picard*. (PhD). Indiana University, United States. Retrieved from <https://search.proquest.com/docview/1628096655?accountid=14701> ProQuest Dissertations & Theses Global database. (3645208)
- Embarki, M., Ouni, S., & Salam, F. (2011). *Speech clarity and coarticulation in Modern Standard Arabic and dialectal Arabic*. Paper presented at the The 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong, China. http://icphs2011.hk.it.cityu.edu.hk/ICPHS_CongressProceedings.htm
- Embarki, M., Ouni, S., Yeou, M., Guilleminot, C., & Al-Maqtari, S. (2011). Acoustic and electromagnetic articulographic study of pharyngealisation: Coarticulatory effects as an index of stylistic and regional variation in Arabic. In Z. M. Hassan & B. Heselwood (Eds.), *Instrumental Studies in Arabic Phonetics* (pp. 193-215). Philadelphia: John Benjamins Publishing.
- Embarki, M., Yeou, M., Guilleminot, C., & Al-Maqtari, S. (2011). *An acoustic study of coarticulation in Modern Standard Arabic and dialectal Arabic: Pharyngealized vs. non-pharyngealized articulation*. Paper presented at the XVI International Congress of Phonetic Sciences (ICPhS), Saarbrücken, Germany. <http://www.icphs2007.de/>
- Esling, J. H. (1996). Pharyngeal consonants and the aryepiglottic sphincter. *Journal of the International Phonetic Association*, 26(2), 65-88. doi:10.1017/S0025100300006125
- Esling, J. H. (1999). The IPA categories “Pharyngeal” and “Epiglottal”: Laryngoscopic observations of pharyngeal articulations and larynx height. *Language and Speech*, 42(4), 349-372. doi:10.1177/00238309990420040101
- Esling, J. H. (2005). There are no back vowels: The laryngeal articulator model. *The Canadian Journal of Linguistics / La revue canadienne de linguistique*, 50(1), 13-44. doi:10.1353/cjl.2007.0007

- Farnetani, E. (1986). *A pilot study of the articulation of /n/ in Italian using electropalatography and airflow measurements*. Paper presented at the 15e Journées d'Etudes sur la Parole, Aix-en-Provence, France.
- Farnetani, E. (1990). V-C-V lingual coarticulation and its spatio-temporal domain. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 93-130). Dordrecht, Netherlands: Kluwer Academic Publishing.
- Farnetani, E., & Recasens, D. (1999). Coarticulation models in recent speech production theories. In W. J. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, Data and Techniques* (pp. 31-65). Cambridge: Cambridge University Press.
- Farnetani, E., & Recasens, D. (2010). Coarticulation and connected speech processes. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The Handbook of Phonetic Sciences, Second Edition*. New Jersey: Blackwell Publishing Ltd.
- Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology*, 18(1), 7-44.
- Fletcher, J., & Harrington, J. (1999). Lip and jaw coarticulation. In W. J. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, Data, and Techniques* (pp. 164-179). Cambridge, New York: Cambridge University Press.
- Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 8(4), 113-133.
- Fowler, C. A. (1981). A relationship between coarticulation and compensatory shortening. *Phonetica*, 38, 35-50.
- Fowler, C. A. (1985). Current perspective on language and speech production: A critical review. In R. Daniloff (Ed.), *Speech Science* (pp. 193-278). London: Taylor and Francis.
- Ghazeli, S. (1977). *Back consonants and backing coarticulation in Arabic*. (PhD). The University of Texas at Austin, United States. Retrieved from <https://search-proquest-com.proxy.bib.uottawa.ca/docview/302856979?accountid=14701> ProQuest Dissertations & Theses Global database. (302856979)
- Gick, B., & Wilson, I. (2006). Excrescent schwa and vowel laxing: Cross-linguistic responses to conflicting articulatory targets. In L. M. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Papers in Laboratory Phonology 8* (pp. 635-659). Berlin: Mouton de Gruyter.
- Gick, B., Wilson, I., & Derrick, D. (2013). *Articulatory Phonetics*. Malden, MA, USA: Wiley-Blackwell.
- Gu, C. (2002). *Smoothing spline ANOVA models*. New York: Springer-Verlag.
- Halle, M. (1995). Feature geometry and feature spreading. *Linguistic Inquiry*, 26(1), 1-46. Retrieved from <http://www.jstor.org/stable/4178887>
- Halle, M., Vaux, B., & Wolfe, A. (2000). On feature spreading and the representation of place of articulation. *Linguistic Inquiry*, 31(3), 387-444. doi:10.1162/002438900554398
- Hammarberg, R. E. (1976). The metaphysics of coarticulation. *Journal of Phonetics*, 4, 353-363.
- Han, J. I. (2007). *The role of vowel contrast in language-specific patterns of vowel-to-vowel coarticulation: Evidence from Korean and Japanese*. Paper presented at the XVI International Congress of Phonetic Sciences (ICPhS), Saarbrücken, Germany.
- Hassan, Z. M., & Esling, J. H. (2007). *Laryngoscopic (articulatory) and acoustic evidence of a prevailing emphatic feature over the word in Arabic*. Paper presented at the XVI International Congress of Phonetic Sciences (ICPhS), Saarbrücken, Germany.
- Hassan, Z. M., & Esling, J. H. (2011). Investigating the emphatic feature in Iraqi Arabic: Acoustic and articulatory evidence of coarticulation. In Z. M. Hassan & B. Heselwood (Eds.),

- Instrumental Studies in Arabic Phonetics* (pp. 217-234). Philadelphia: John Benjamins Publishing.
- Hassan, Z. M., Esling, J. H., Moisik, S. R., & Crevier-Buchman, L. (2011). *Aryepiglottic trilled variants of /ʕ ħ/ in Iraqi Arabic*. Paper presented at the The 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong, China. http://icphs2011.hk.lt.cityu.edu.hk/ICPHS_CongressProceedings.htm
- Henke, W. L. (1966). *Dynamic articulatory model of speech production using computer simulation*. (Doctor of Philosophy). Massachusetts Institute of Technology, Massachusetts, United States. Retrieved from <https://search.proquest.com/docview/302308985?accountid=14701> ProQuest Dissertations & Theses Global database. (302308985)
- Herzallah, R. S. (1990). *Aspects of Palestinian Arabic phonology: A Nonlinear Approach*. (Doctorate of Philosophy). Cornell University, United States. ProQuest Dissertations & Theses Global database. (303851336)
- Heselwood, B. (2007). The ‘tight approximant’ variant of the Arabic ‘ayn’. *Journal of the International Phonetic Association*, 37, 1-32.
- Heselwood, B., & Al-Tamimi, F. (2011). A study of the laryngeal and pharyngeal consonants in Jordanian Arabic using nasoendoscopy, videofluoroscopy and spectrography. In Z. M. Hassan & B. Heselwood (Eds.), *Instrumental Studies in Arabic Phonetics* (pp. 101-127). Philadelphia: John Benjamins Publishing.
- Hixon, T. J., Weismer, G., & Hoit, J. D. (2014). *Preclinical Speech Science : Anatomy, Physiology, Acoustics, Perception (Second Edition)*. San Diego, CA: Plural Publication.
- Holes, C. (1983). Bahraini dialects: sectarian differences and the sedentary/nomadic split. *Zeitschrift für Arabische Linguistik*, 10, 7-38. Retrieved from <https://www.jstor.org/stable/43530422>
- Holes, C. (1990). *Gulf Arabic*. London: Routledge.
- Hueber, T., Chollet, G., Denby, B., & Stone, M. (2008). Acquisition of ultrasound, video and acoustic speech data for a silent-speech interface application. *Proceedings of the 8th International Seminar on Speech Production*, 365-369.
- Ingham, B. (1982). *North East Arabian Dialects*. London: Routledge.
- Iskarous, K., & Kavitskaya, D. (2010). The Interaction between Contrast, Prosody, and Coarticulation in Structuring Phonetic Variability. *Journal of Phonetics*, 38(4), 625-639. doi:10.1016/j.wocn.2010.09.004
- Jaber, A., Omari, O., & Al-Jarrah, R. (In press). The domain of emphasis spread in Arabic: Evidence from Urban Jordanian Arabic. *Lingua*. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0024384118306545>
- Jakobson, R., Fant, G., & Halle, M. (1951). *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates*. Cambridge, MA: M.I.T. Press.
- Jongman, A., Herd, W., Al-Masri, M., Sereno, J., & Combest, S. (2011). Acoustics and perception of emphasis in Urban Jordanian Arabic. *Journal of Phonetics*, 39(1), 85-95. doi:10.1016/j.wocn.2010.11.007
- Kahn, M. (1975). Arabic emphatics: The evidence for cultural determinants of phonetic sex-typing. *Phonetica*, 31, 38-50. doi:10.1159/000259648
- Kavitskaya, D. (1997). *Aerodynamic Constraints on the Production of Palatalized Trills: The Case of the Slavic Trilled [r]*. Paper presented at the 5th European Conference on Speech

- Communication and Technology, Rhodes, Greece. https://www.isca-speech.org/archive/eurospeech_1997/e97_0751.html
- Keating, P. A. (1985). Universal phonetics and the organization of grammars. In V. Fromkin (Ed.), *Phonetic Linguistics: Essays in Honor of Peter Ladefoged* (pp. 115-132). Orlando, FL: Academic Press.
- Keating, P. A. (1988a). Underspecification in phonetics. *Phonology*, 5(02), 275-292. doi:10.1017/s095267570000230x
- Keating, P. A. (1988b). The window model of coarticulation: Articulatory evidence. *UCLA Working Papers in Phonetics*, 69, 3-29.
- Keating, P. A. (1990). The window model of coarticulation: Articulatory evidence. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech* (pp. 451-470). Cambridge: Cambridge University Press.
- Kent, R. D. (1998). *The Speech Sciences*. San Diego, California: Singular Publishing Group.
- Kent, R. D., Carney, P. J., & Severeid, L. R. (1974). Velar movement and timing: Evaluation of a model of binary control. *Journal of Speech and Hearing Research*, 17, 57-89.
- Kharlamov, V. (2012). *Incomplete neutralization and task effects in experimentally-elicited speech: Evidence from the production and perception of word-final devoicing in Russian*. (Doctorate of Philosophy). University of Ottawa, Canada. Retrieved from <http://hdl.handle.net/10393/22809> ProQuest Dissertations & Theses Global database. (1356256712)
- Khattab, G., Al-Tamimi, F., & Heselwood, B. (2006). Acoustic and auditory differences in the /t-/ /t/ opposition in male and female speakers of Jordanian Arabic. In S. Boudelaa (Ed.), *Perspectives on Arabic Linguistics XVI: Papers from the Sixteenth Annual Symposium on Arabic Linguistics* (pp. 131-160). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Kier, W. M., & Smith, K. K. (1985). Tongues, tentacles and trunks: The biomechanics of movement in muscular-hydrostats. *Zoological Journal of the Linnean Society*, 83, 307-324. Retrieved from <https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1096-3642.1985.tb01178.x>
- Kozhevnikov, V. A., & Chistovich, L. A. (1965). *Speech: Articulation and Perception*. Washington: Joint Publications Research Service.
- Kuehn, D. P., & Azzam, N. A. (1978). Anatomical Characteristics of Palatoglossus and the Anterior Faucial Pillar. *The Cleft Palate Journal*, 15(4). Retrieved from <http://cleftpalatejournal.pitt.edu/ojs/cleftpalate/article/view/749>
- Kuriyagawa, F., Sawashima, M., Niimi, S., & H., H. (1988). Electromyographic study of emphatic consonants in standard Jordanian Arabic. *Folia Phoniatrica et Logopaedica*, 40(3), 117-122. doi:10.1159/000265897
- Lardi, W. (1983). *Pharyngealization in Libyan (Tripoli) Arabic : An instrumental study*. (PhD). University of Eginburgh, Retrieved from <http://hdl.handle.net/1842/6717> (346529)
- Laufer, A., & Baer, T. (1988). The emphatic and pharyngeal sounds in Hebrew and Arabic. *Language and Speech*, 31(2), 181-205. doi:10.1177/002383098803100205
- Laufer, A., & Condax, I. D. (1978). The epiglottis as an articulator. *The Journal of the Acoustical Society of America*, 64(S1). doi:10.1121/1.2004454
- Laufer, A., & Condax, I. D. (1981). The function of the epiglottis in speech. *Language & Speech*, 24(1), 39-62. doi:10.1177/002383098102400103

- Lavoie, L. M., & Cohn, A. C. (1999). *Sesquisyllables of English: The structure of vowel-liquid syllables*. Paper presented at the The XIVth International Congress of Phonetic Sciences, San Francisco, CA.
- Lawson, E., Stuart-Smith, J., Scobbie, J. M., & Nakai, S. (Producer). (2018). Seeing Speech: An articulatory web resource for the study of Phonetics.
- Lehn, W. (1963). Emphasis in Cairo Arabic. *Language*, 39(1), 29-39. doi:10.2307/410760
- Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, 48(4), 839-862. doi:10.2307/411991
- Lindblom, B. (1986). Phonetic universals in vowel systems. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental Phonology* (pp. 13-44). Orlando: Academic Press.
- Lindblom, B., Pauli, S., & Sundberg, J. (1975). Modelling coarticulation in apical stops. In G. Fant (Ed.), *Speech Communication II* (pp. 87-94). Uppsala: Almqvist and Wiksell.
- MacNeilage, P. (1970). Motor control and serial ordering of speech. *Psychological Review*, 77, 182-196.
- Maddieson, I. (1984). *Patterns of Sounds*. New York: Cambridge University Press.
- Maeda, S. (1993). Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In M. K. Huffman & R. A. Krakow (Eds.), *Phonetics and Phonology: Nasals, Nasalization, and the Velum* (Vol. 5, pp. 147-167). San Diego: Academic Press.
- Magen, H. (1984). Vowel-to-vowel coarticulation in English and Japanese. *The Journal of the Acoustical Society of America*, 75(S1).
- Magen, H. (1997). The extent of vowel-to-vowel coarticulation in English. *Journal of Phonetics*, 25(2), 187-205. Retrieved from <https://doi.org/10.1006/jpho.1996.0041>
- Manuel, S. Y. (1987). Output constraints and cross-language differences in coarticulation. *The Journal of the Acoustical Society of America*, 82(S1), S115-S115. doi:10.1121/1.2024600
- Manuel, S. Y. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *The Journal of the Acoustical Society of America*, 88, 1286-1298. Retrieved from <https://doi.org/10.1121/1.399705>
- Manuel, S. Y. (1999). Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts. In W. J. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, Data and Techniques* (pp. 179-200). Cambridge: Cambridge University Press.
- Manuel, S. Y., & Krakow, R. A. (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins Laboratories Status Report*, 77/78, 69-78. Retrieved from <http://www.haskins.yale.edu/sr/sr077/sr077.html>
- McAuliffe, M., Socolof, M., Stengel-Eskin, E., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner [Computer program] (Version 1.0.0). Retrieved from <http://montrealcorpus-tools.github.io/Montreal-Forced-Aligner/>
- McCarthy, J. J. (1991). Semitic gutturals and distinctive feature theory. In B. Comrie & M. Eid (Eds.), *Perspectives on Arabic Linguistics III* (pp. 63-91). Amsterdam: John Benjamins Publishing.
- McCarthy, J. J. (1994). The phonetics and phonology of Semitic pharyngeals. In P. A. Keating (Ed.), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*. Cambridge: Cambridge University Press.
- McOmber, M. L. (1996). Phonemic pharyngealization. In M. Eid (Ed.), *Perspectives on Arabic Linguistics VIII: Papers from the Eighth Annual Symposium on Arabic Linguistics* (Vol. 134, pp. 233-258). Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Merrill, J. (2014). *A typological overview of consonant mutation*. Retrieved from <http://linguistics.berkeley.edu/~merrill/>
- Mielke, J. (2015). An ultrasound study of Canadian French rhotic vowels with polar smoothing spline comparisons. *The Journal of the Acoustical Society of America*, 137, 2858-2869. Retrieved from <http://dx.doi.org/10.1121/1.4919346>
- Mielke, J., Carignan, C., & Thomas, E. R. (2017). The articulatory dynamics of pre-velar and pre-nasal /ae/-raising in English: An ultrasound study. *J Acoust Soc Am*, 142(1), 332. doi:10.1121/1.4991348
- Mok, P. K. (2012). Does vowel inventory density affect vowel-to-vowel coarticulation? *Language and Speech*, 56(2), 191-209. doi:10.1177/0023830912443948
- Mok, P. K., & Hawkins, S. (2004). *Effects of phonemic vs allophonic density and stress on vowel-to-vowel coarticulation in Cantonese and Beijing Mandarin*. Paper presented at the International Symposium on Chinese Spoken Language Processing, Hong Kong, China.
- Moll, K. L., & Daniloff, R. G. (1971). Investigation of the Timing of Velar Movements during Speech. *The Journal of the Acoustical Society of America*, 50, 678-684. doi:<https://doi.org/10.1121/1.1912683>
- Nicenboim, B., Roettger, T. B., & Vasishth, S. (2018). Using meta-analysis for evidence synthesis: The case of incomplete neutralization in German. *Journal of Phonetics*, 70, 39-55. doi:10.1016/j.wocn.2018.06.001
- Ohala, J. J. (1971). Monitoring soft palate movements in speech. *The Journal of the Acoustical Society of America*, 50(1A), 140. doi:10.1121/1.1977664
- Ohala, J. J. (1975). Phonetic explanations for nasal sound patterns. In C. A. Ferguson, L. M. Hyman, & J. J. Ohala (Eds.), *Nasálfest: Papers from a Symposium on Nasals and Nasalization* (pp. 289-316). Stanford, California: Language Universals Project.
- Ohala, J. J. (1990). The generality of articulatory binding: Comments on Kingston's paper. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech* (pp. 435-443). England, New York: Cambridge University Press.
- Öhman. (1966). Coarticulation in VCV Utterances: Spectrographic Measurements. *The Journal of the Acoustical Society of America*, 39(1), 151-168. doi:10.1121/1.1909864
- Öhman. (1967). Numerical Model of Coarticulation. *The Journal of the Acoustical Society of America*, 41(2), 310-320. doi:10.1121/1.1910340
- Orgun, C. O. (2001). English r-insertion in optimality theory. In *Natural Language and Linguistic Theory* (pp. 737-749). Dordrecht, Holland: D. Reidel Publishing.
- Recasens, D. (1989). Long range coarticulation effects for tongue dorsum contact in VCVCV sequences. *Speech Communication*, 8, 293-307. doi:[https://doi.org/10.1016/0167-6393\(89\)90012-5](https://doi.org/10.1016/0167-6393(89)90012-5)
- Recasens, D. (2012). A study of jaw coarticulatory resistance and aggressiveness for Catalan consonants and vowels. *J Acoust Soc Am*, 132(1), 412-420. doi:10.1121/1.4726048
- Recasens, D. (2014). *Coarticulation and Sound Change in Romance* (Vol. 329). Amsterdam: John Benjamins Publishing Company.
- Recasens, D. (2015). The effect of stress and speech rate on vowel coarticulation in Catalan vowel-consonant-vowel sequences. *Journal of Speech and Hearing Research*, 58(5), 1407-1424. doi:10.1044/2015_JSLHR-S-14-0196

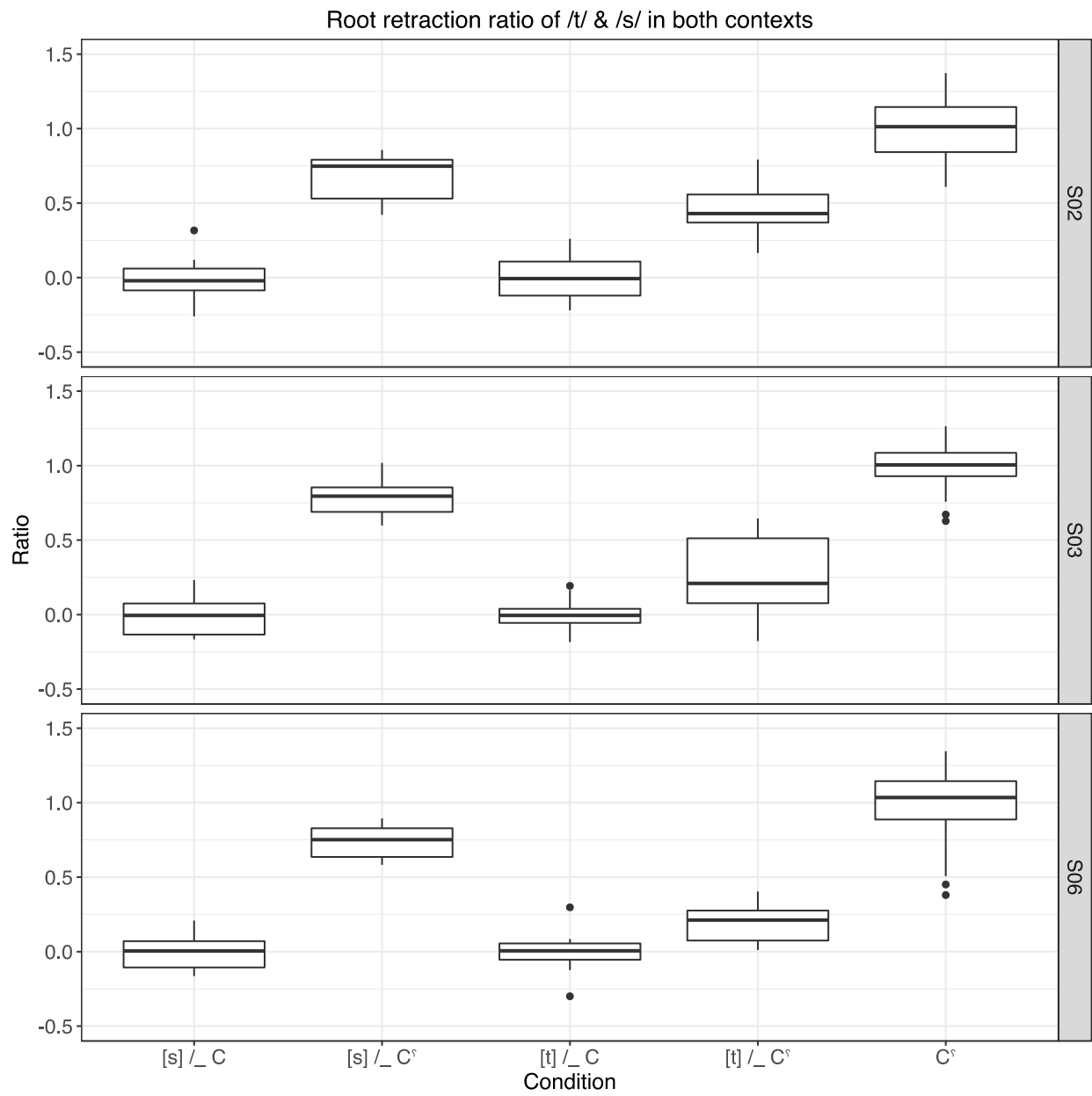
- Recasens, D., & Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *J Acoust Soc Am*, 125(4), 2288-2298. doi:10.1121/1.3089222
- Recasens, D., & Mira, M. (2015). Articulatory reduction and coarticulation in Catalan three-consonant sequences. *The Journal of the Acoustical Society of America*, 137(1), 397-406. doi:10.1121/1.4904499
- Recasens, D., & Pallarès, M. D. (1999). A study of /j/ and /r/ in the light of the “DAC” coarticulation model. *Journal of Phonetics*, 27, 143-169. Retrieved from <https://www-sciencedirect-com.proxy.bib.uottawa.ca/science/article/pii/S009544709990092X>
- Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *The Journal of the Acoustical Society of America*, 102(1), 544-561. doi:10.1121/1.419727
- Recasens, D., & Rodríguez, C. (2016). A study on coarticulatory resistance and aggressiveness for front lingual consonants and vowels using ultrasound. *Journal of Phonetics*, 59, 58-75. doi:10.1016/j.wocn.2016.09.002
- Recasens, D., & Rodríguez, C. (2018). Contextual and syllabic effects in heterosyllabic consonant sequences. An ultrasound study. *Speech Communication*, 96, 150-167. doi:10.1016/j.specom.2017.11.009
- Roettger, T. B., Winter, B., Grawunder, S., Kirby, J., & Grice, M. (2014). Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics*, 43, 11-25. doi:10.1016/j.wocn.2014.01.002
- Rose, S. (1996). Variable laryngeals and vowel lowering. *Phonology*, 13(1), 73-117. doi:10.1017/S0952675700000191
- Sanders, I., & Mu, L. (2013). A three-dimensional atlas of human tongue muscles. *Anat Rec (Hoboken)*, 296(7), 1102-1114. doi:10.1002/ar.22711
- Scarborough, R., Zellou, G., Mirzayan, A., & Rood, D. S. (2015). Phonetic and phonological patterns of nasality in Lakota vowels. *Journal of the International Phonetic Association*, 45(03), 289-309. doi:10.1017/s0025100315000171
- Schwartz, J. L., Boë, L. J., Vallée, N., & Abry, C. (1997). Major trends in vowel system inventories. *Journal of Phonetics*, 25(3), 233-253. doi:10.1006/jpho.1997.0044
- Scobbie, J. M., & Stuart-Smith, J. (2008). Quasi-phonemic contrast and the fuzzy inventory: Examples from Scottish English. In *Contrast, Perception and Acquisition: Selected Papers from the Second International Conference on Contrast in Phonology* (pp. 87-113). Berlin: Mouton.
- Scobbie, J. M., Wrench, A. A., & van der Linden, M. (2008). Head-probe stabilization in ultrasound tongue imaging using a headset to permit natural head movement. *Proceedings of the 8th International Seminar on Speech Production*, 373-376.
- Seikel, J. A., King, D. W., & Drumright, D. G. (2000). *Anatomy and Physiology for Speech Language and Hearing (Second edition)*. San Diego, California: Singular Publishing Group.
- Shahin, K. (1996). Accessing pharyngeal place in Palestinian Arabic. In M. Eid & D. Parkinson (Eds.), *Perspectives on Arabic Linguistics IX: Papers from the Ninth Annual Symposium on Arabic Linguistics* (Vol. 141, pp. 131-150). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Shahin, K. (1997). Acoustics of pharyngealization vs. uvularization harmony. In M. Eid & R. R. Ratcliffe (Eds.), *Perspectives on Arabic Linguistics X: Papers from the Tenth Annual*

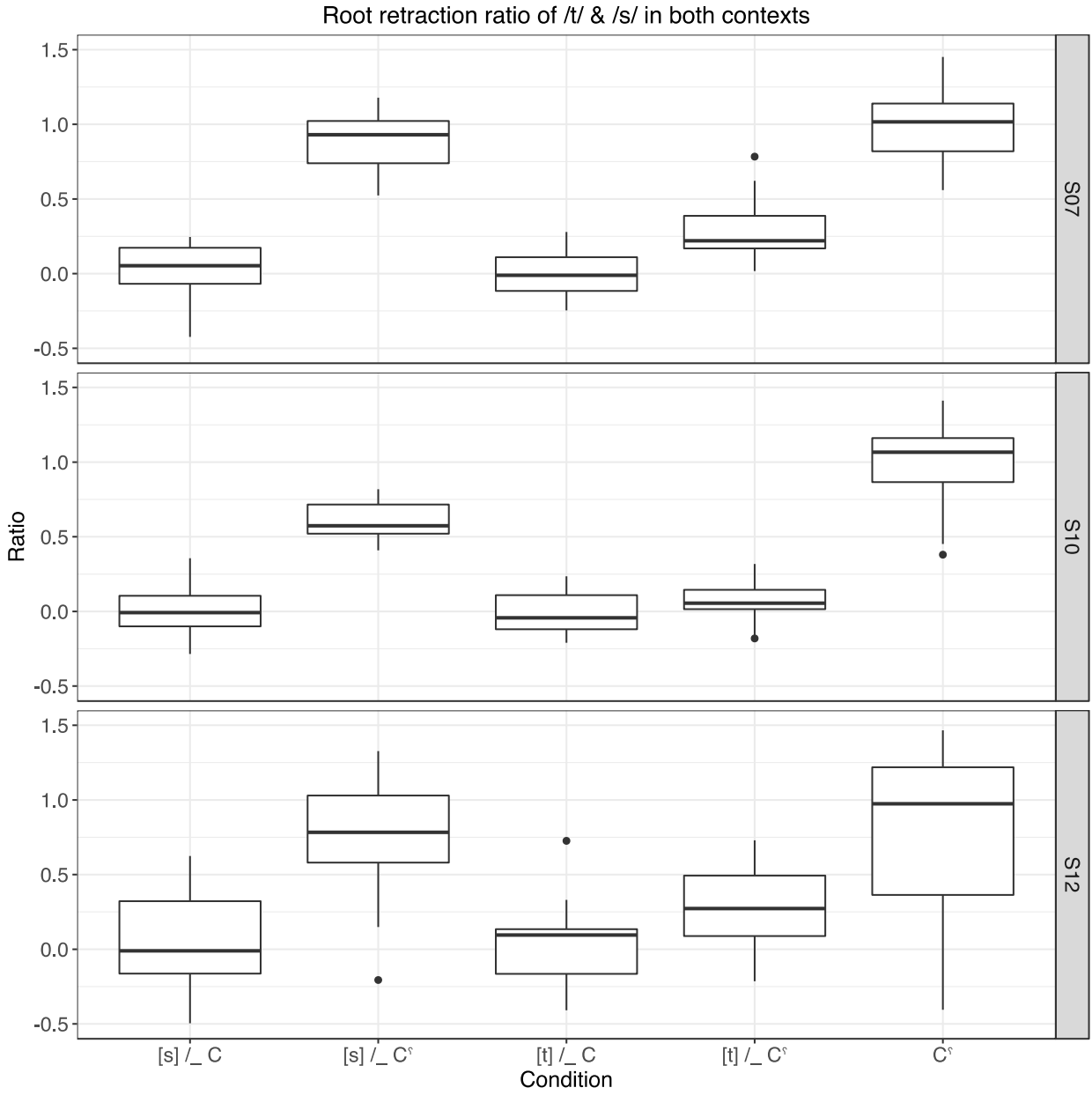
- Symposium on Arabic Linguistics* (Vol. 153, pp. 215-238). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Shahin, K. (2011). A phonetic study of guttural laryngeals in Palestinian Arabic using laryngoscopic and acoustic analysis. In Z. M. Hassan & B. Heselwood (Eds.), *Instrumental Studies in Arabic Phonetics* (pp. 129-140). Philadelphia: John Benjamins Publishing.
- Shar, S., & Ingram, J. (2011). *MRI investigation of Arabic gutturals*. Paper presented at the The 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong, China. http://icphs2011.hk.lt.cityu.edu.hk/ICPHS_CongressProceedings.htm
- Shevelov, G. Y. (1979). *A historical phonology of the Ukrainian language*. Heidelberg: C. Winter Universitätsverlag.
- Solé, M. J. (1992). Phonetic and phonological processes: The case of nasalization. *Language and Speech*, 35(1/2), 29-43. doi:10.1177/002383099203500204
- Solé, M. J. (1995). Spatio-temporal patterns of velopharyngeal action in phonetic and phonological nasalization. *Language and Speech*, 38(1), 1-24. doi:10.1177/002383099503800101
- Solé, M. J. (2002). Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics*, 30(4), 655-688. doi:10.1006/jpho.2002.0179
- Spajić, S., Ladefoged, P., & Bhaskararao, P. (1996). The trills of Toda. *Journal of the International Phonetic Association*, 26(1), 1-21. doi:<https://doi-org.proxy.bib.uottawa.ca/10.1017/S0025100300005296>
- Steriade, D. (1987). *Redundant values*. Paper presented at the Chicago Linsuitic Soceity (CLS 23), Chicago, Illinois.
- Stone, M., Woo, J., Lee, J., Poole, T., Seagraves, A., Chung, M., . . . Blemker, S. S. (2018). Structure and variability in human tongue muscle anatomy. *Comput Methods Biomech Biomed Eng Imaging Vis*, 6(5), 499-507. doi:10.1080/21681163.2016.1162752
- Su, L. S., Li, K. P., & Fu, K. S. (1974). Identification of speakers by use of nasal coarticulation. *The Journal of the Acoustical Society of America*, 56(6), 1876-1883. doi:10.1121/1.1903526
- Takano, S., & Honda, K. (2007). An MRI analysis of the extrinsic tongue muscles during vowel production. *Speech Communication*, 49, 49-58.
- Thelwall, R., & Sa'adeddin, M. A. (1999). Arabic. In *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet* (pp. 51-54). Cambridge: Cambridge University Press.
- Wood, S. A. J. (1996). Assimilation or coarticulation? Evidence from the temporal co-ordination of tongue gestures for the palatalization of Bulgarian alveolar stops. *Journal of Phonetics*, 24, 139-164.
- Yeou, M. (1997). Locus equation and the degree of coarticulation in Arabic consonants. *Phonetica*, 54(3-4), 187-202. doi:10.1159/000262221
- Younes, M. A. (1994). On emphasis and /r/ in Arabic. In M. Eid, V. Cantarino, & K. Walters (Eds.), *Perspectives on Arabic Linguistics VI: Papers from the 6th Annual Symposium on Arabic Linguistics* (Vol. 115, pp. 213-235). Amsterdam/Philadelphia: John Benjamins Publishing.
- Zawaydeh, B. A. (1997). An acoustic analysis of uvularization spread in Ammani-Jordanian Arabic. *Studies in the Linguistic Sciences*, 27(1), 185-200. Retrieved from <https://www.ideals.illinois.edu/bitstream/handle/2142/11582/SLS1997v27.1-14Zawaydeh.pdf?sequence=2>

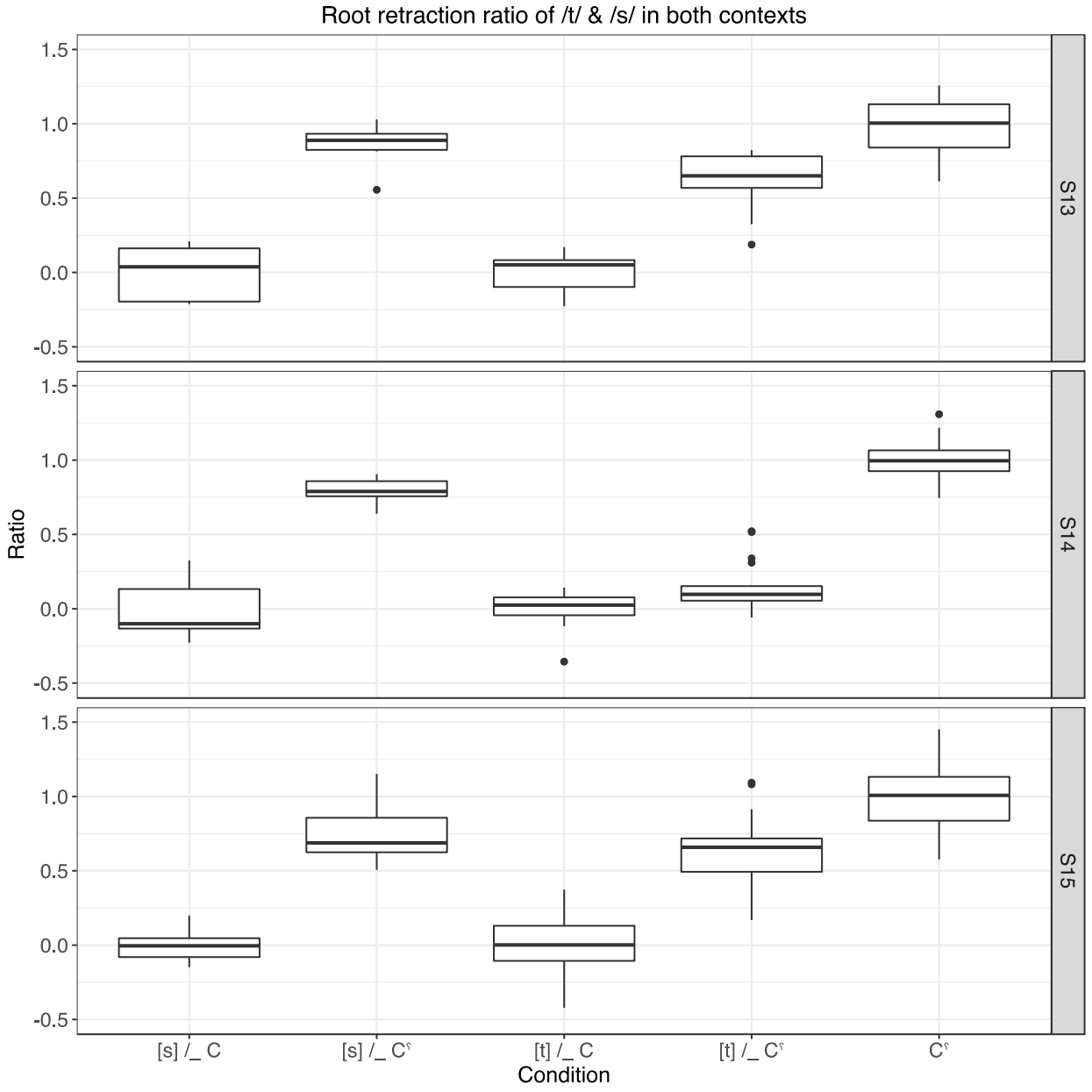
- Zawaydeh, B. A. (1998). Gradient uvularization spread in Ammani-Jordanian Arabic. In E. Benmamoun, M. Eid, & N. Haeri (Eds.), *Perspectives on Arabic Linguistics XI: Papers from the Eleventh Annual Symposium on Arabic Linguistics* (Vol. 167, pp. 117-142). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Zawaydeh, B. A. (1999). *The phonetics and phonology of gutturals in Arabic*. Indiana University, ProQuest Dissertations and Theses database.
- Zawaydeh, B. A. (2003). The interaction of the phonetics and phonology of gutturals. In J. Local, R. Ogden, R. Temple, M. E. Beckman, & J. Kingston (Eds.), *Phonetic interpretation: Papers in laboratory phonology VI* (pp. 279-292). Cambridge: Cambridge University Press.
- Zawaydeh, B. A., & de Jong, K. (2011). The phonetics of localising uvularisation in ammani-jordanian arabic. In Z. M. Hassan & B. Heselwood (Eds.), *Instrumental Studies in Arabic Phonetics* (pp. 257-276). Philadelphia: John Benjamins Publishing.
- Zellou, G. (2013). Consonant harmony in Moroccan Arabic: Similarity and incomplete neutralization. *The Journal of the Acoustical Society of America*, 133, 3572. doi:10.1121/1.4806549
- Zemlin, W. R. (1998). *Speech and Hearing Science: Anatomy and Physiology (Fourth edition)*. Boston, MA: Allyn & Bacon.
- Zeroual, C., & Clements, G. N. (2015). The feature [pharyngeal]. In A. R. Rialland, Rachid; Hulst, Harry van der (Ed.), *Features in Phonology and Phonetics: Posthumous Writings by Nick Clements and Coauthors* (pp. 247-276). Berlin, Boston: DE GRUYTER MOUTON.
- Zeroual, C., Esling, J. H., & Hoole, P. (2011). EMA, endoscopic, ultrasound and acoustic study of two secondary articulations in Moroccan Arabic. In Z. M. Hassan & B. Heselwood (Eds.), *Instrumental Studies in Arabic Phonetics* (pp. 277-297). Philadelphia: John Benjamins Publishing.

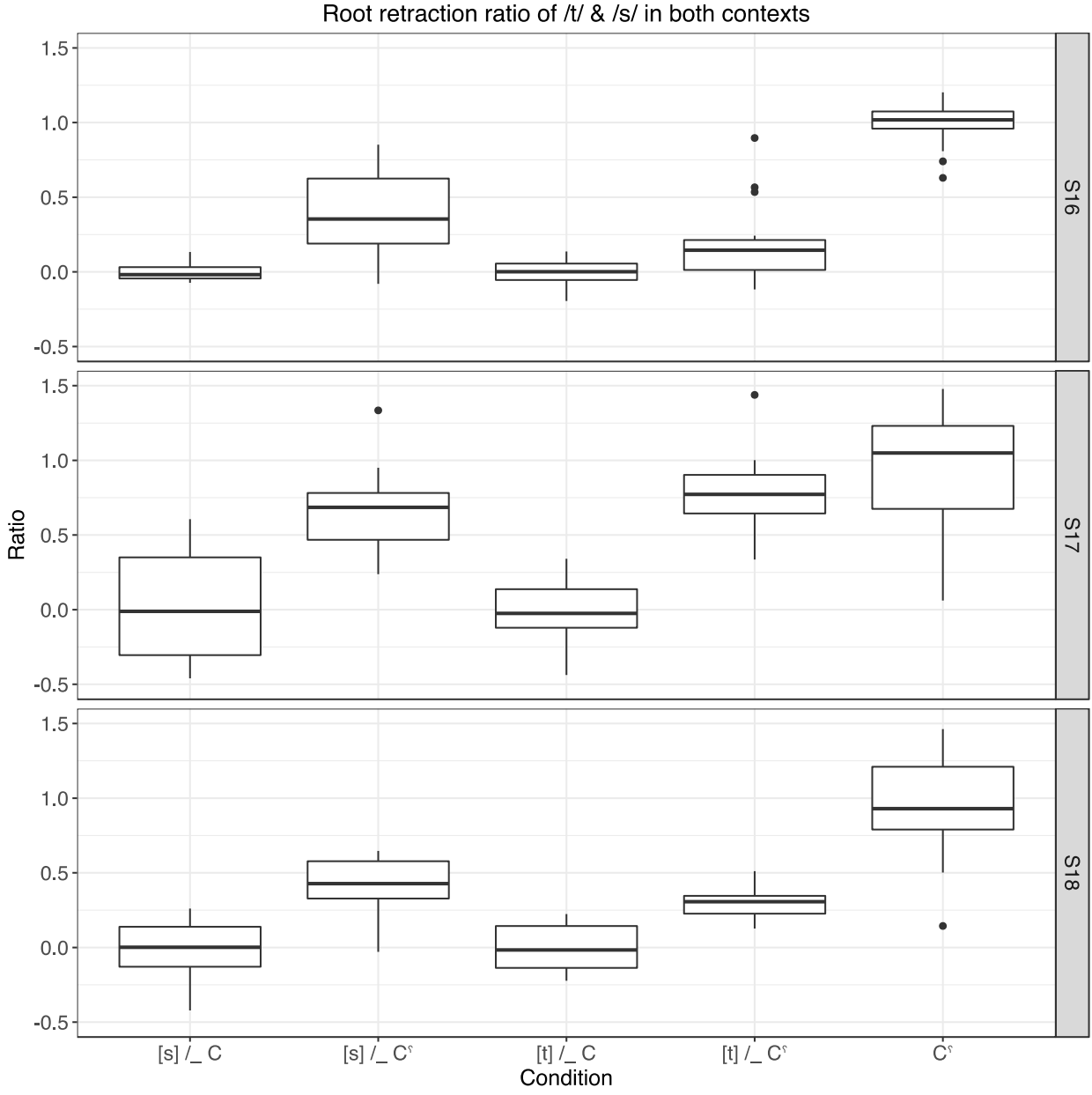
Appendices

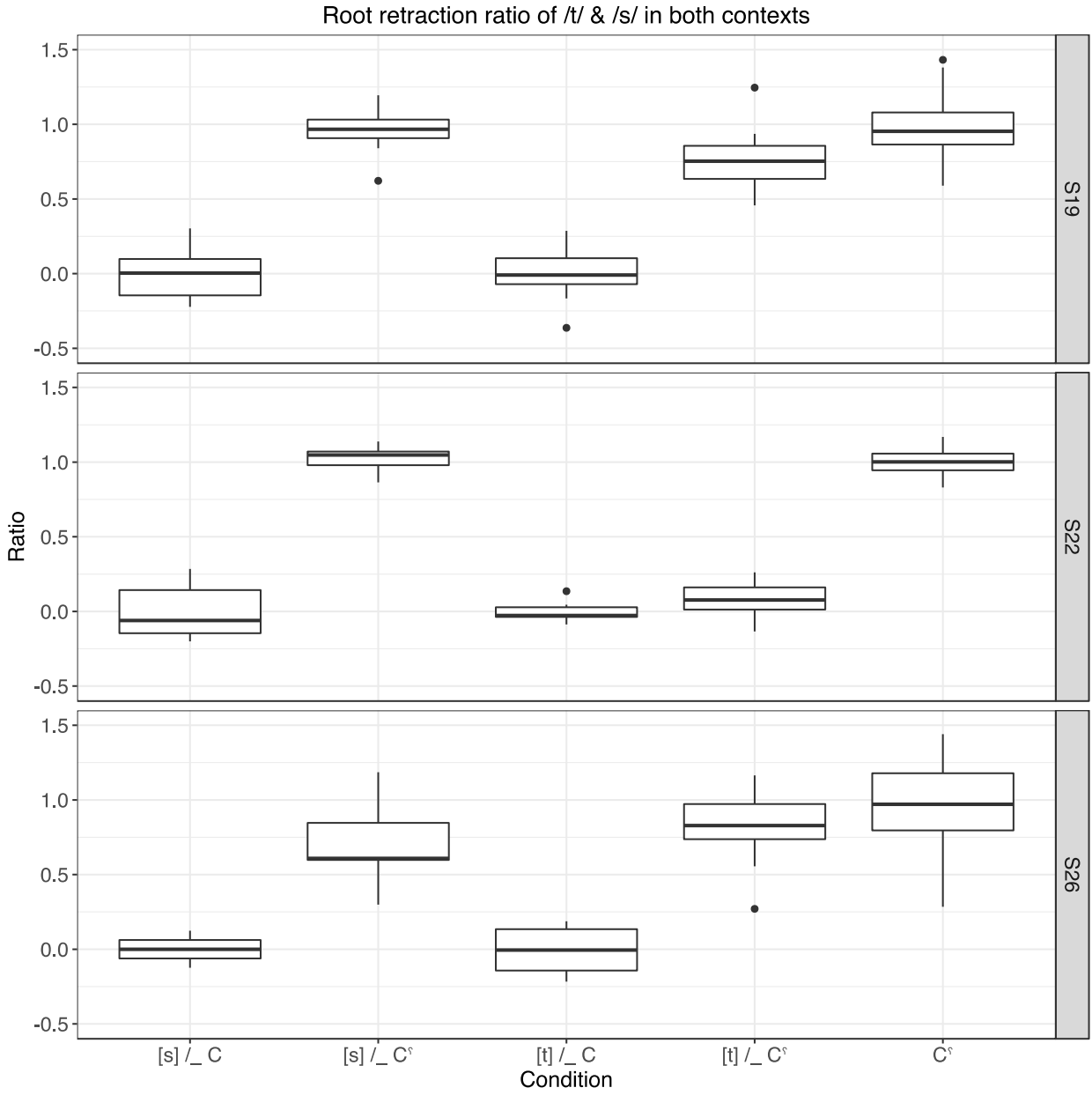
Appendix 1: Scatterplots of the root retraction ratio of /t/ and /s/



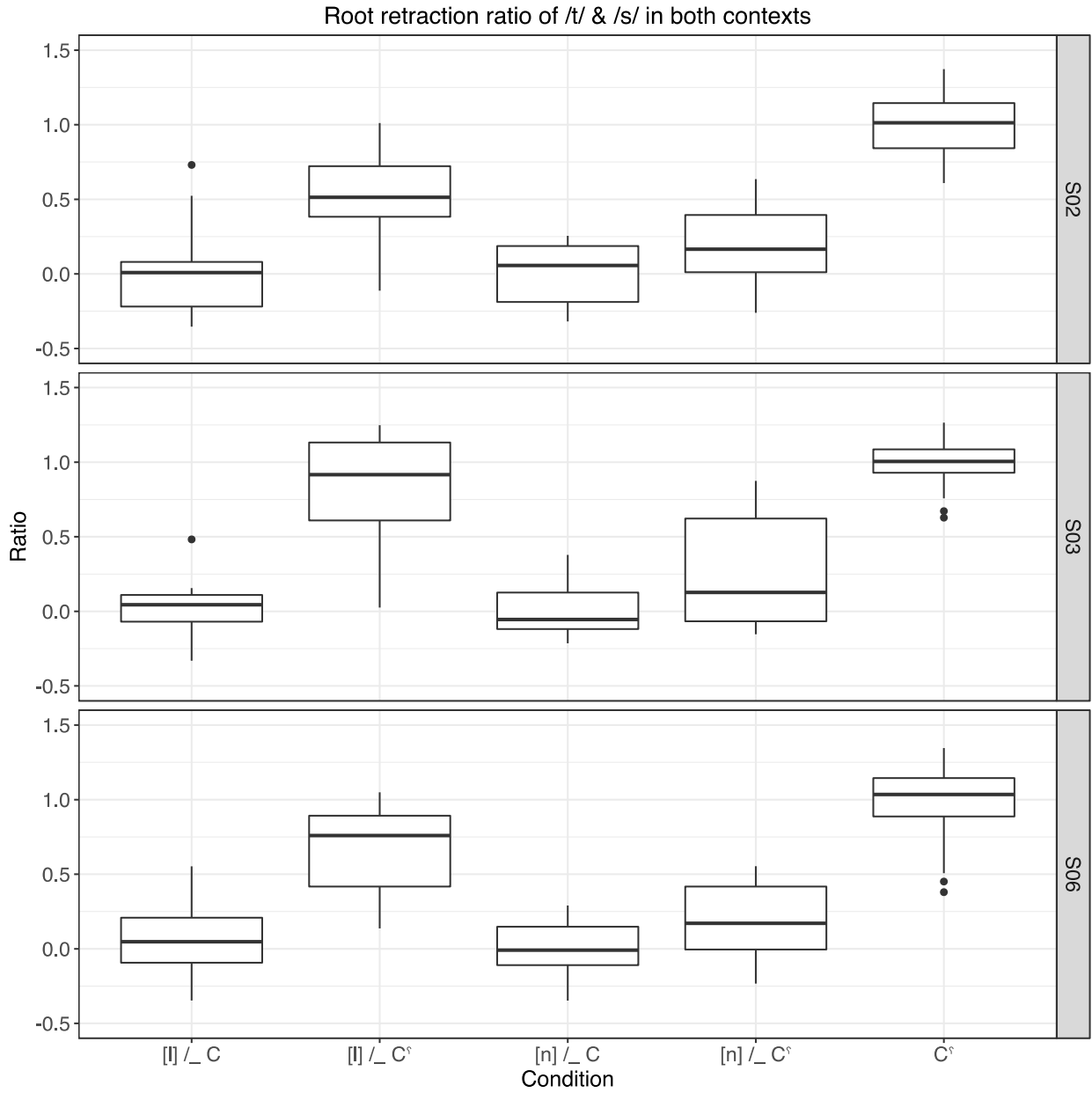


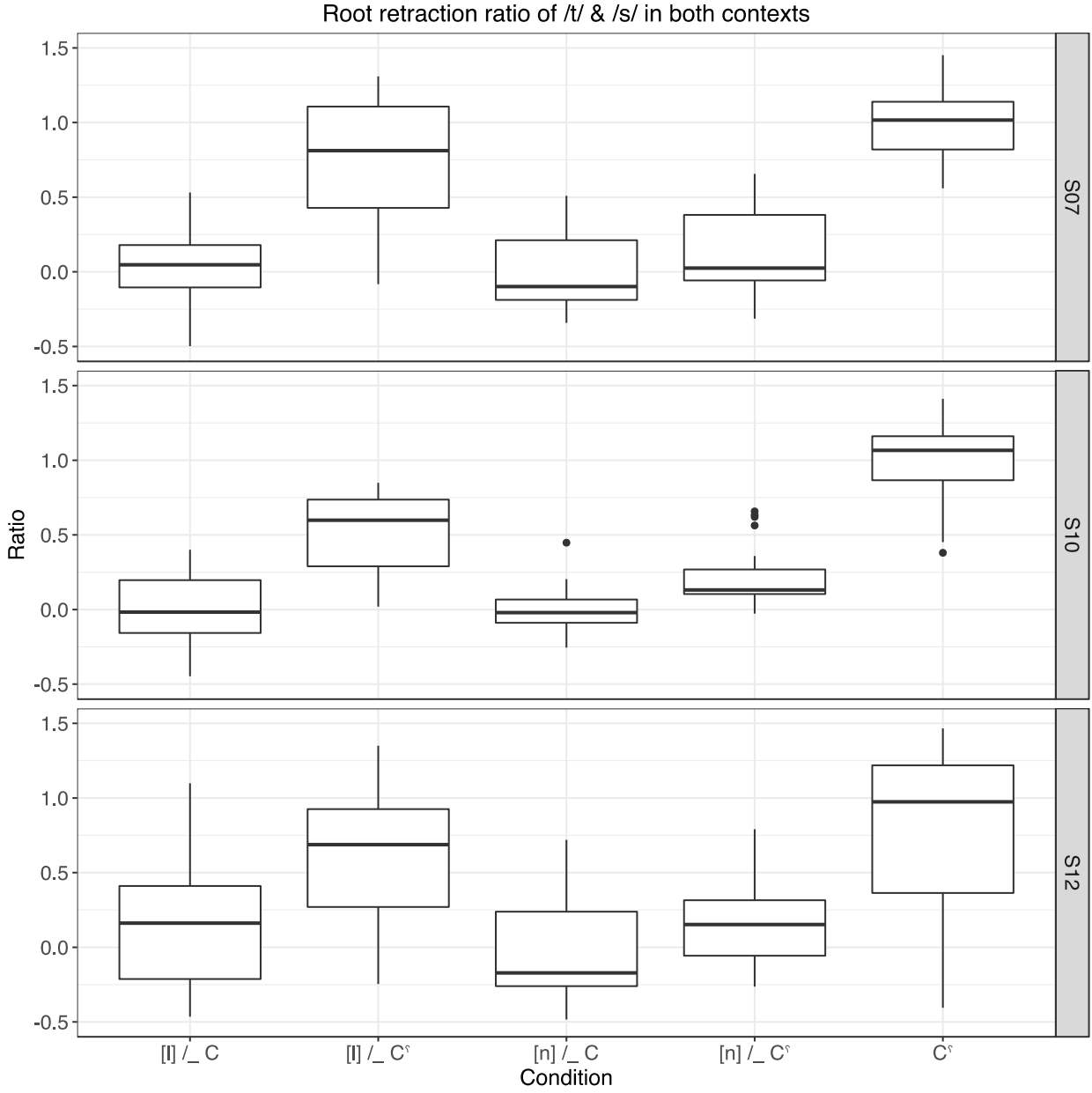


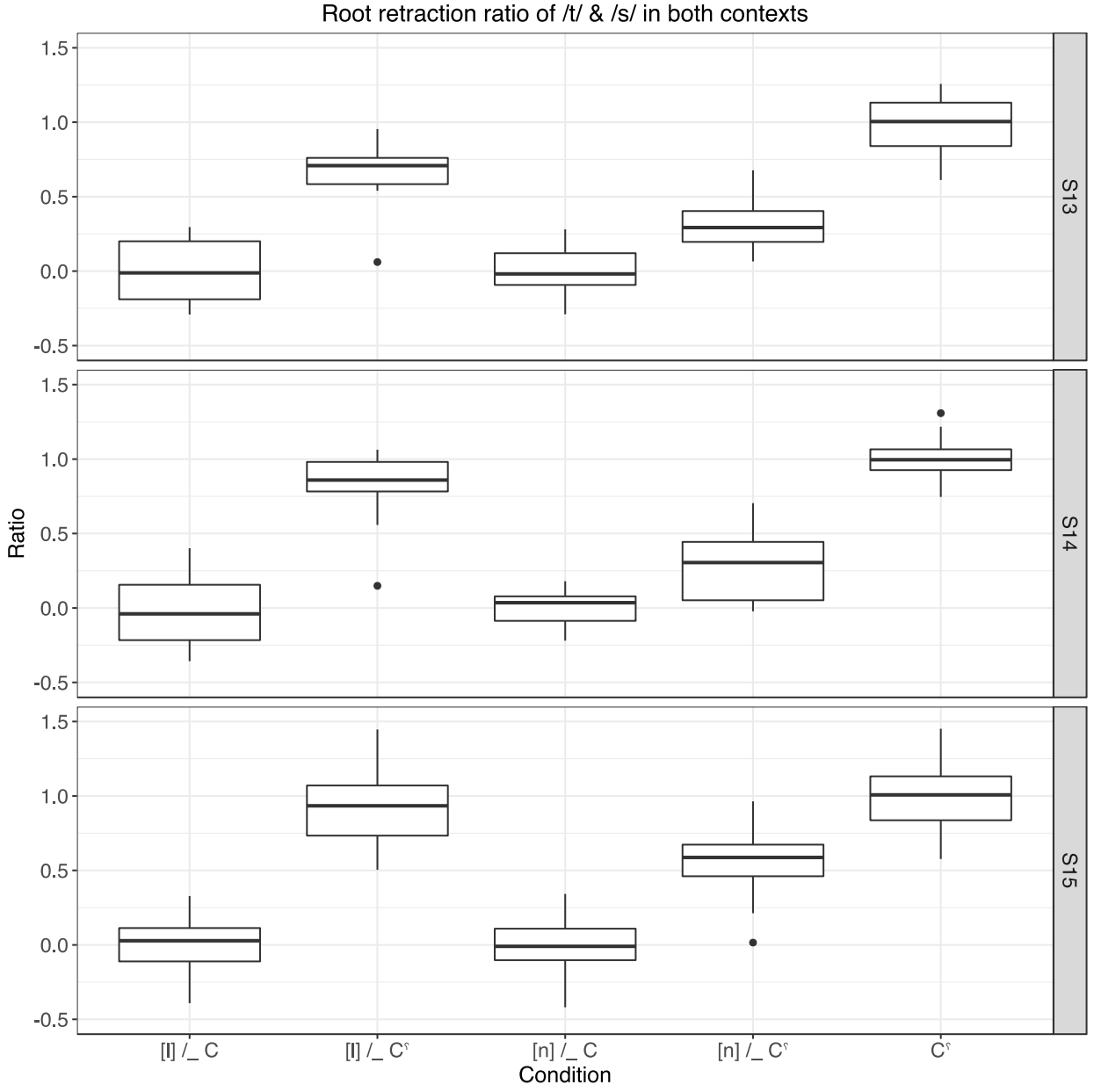


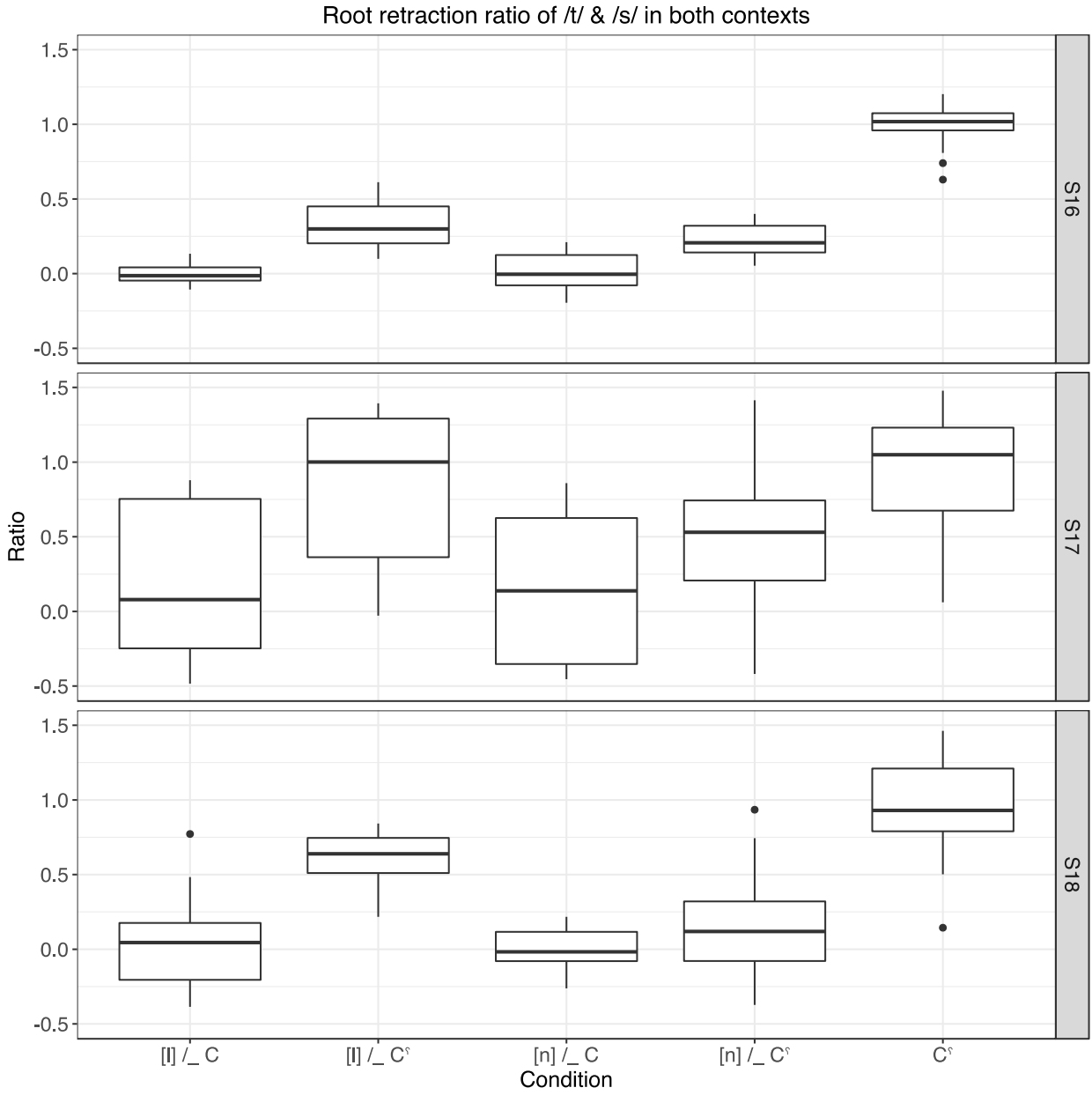


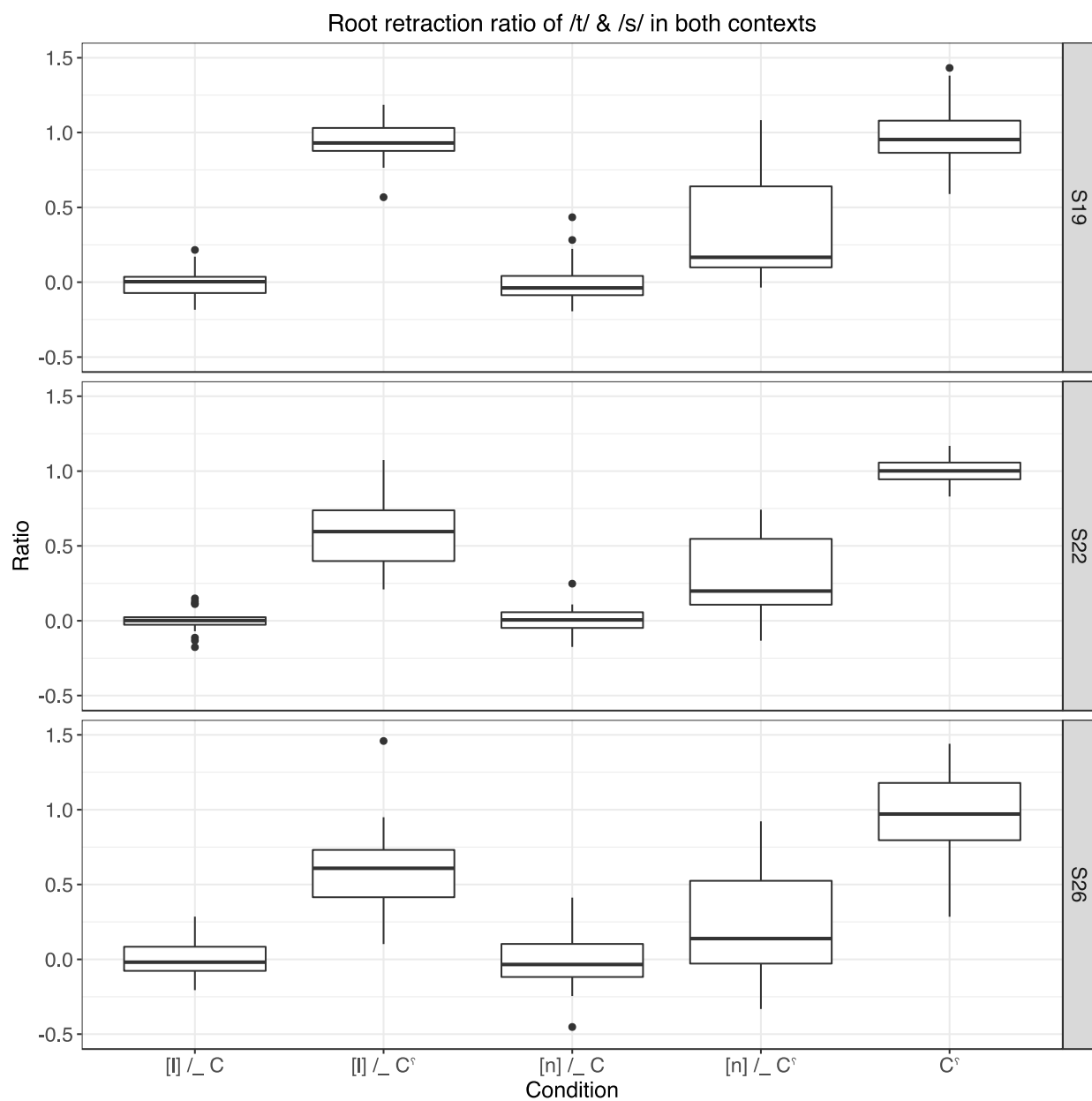
Appendix 2: Scatterplots of the root retraction ratio of /n/ and /l/



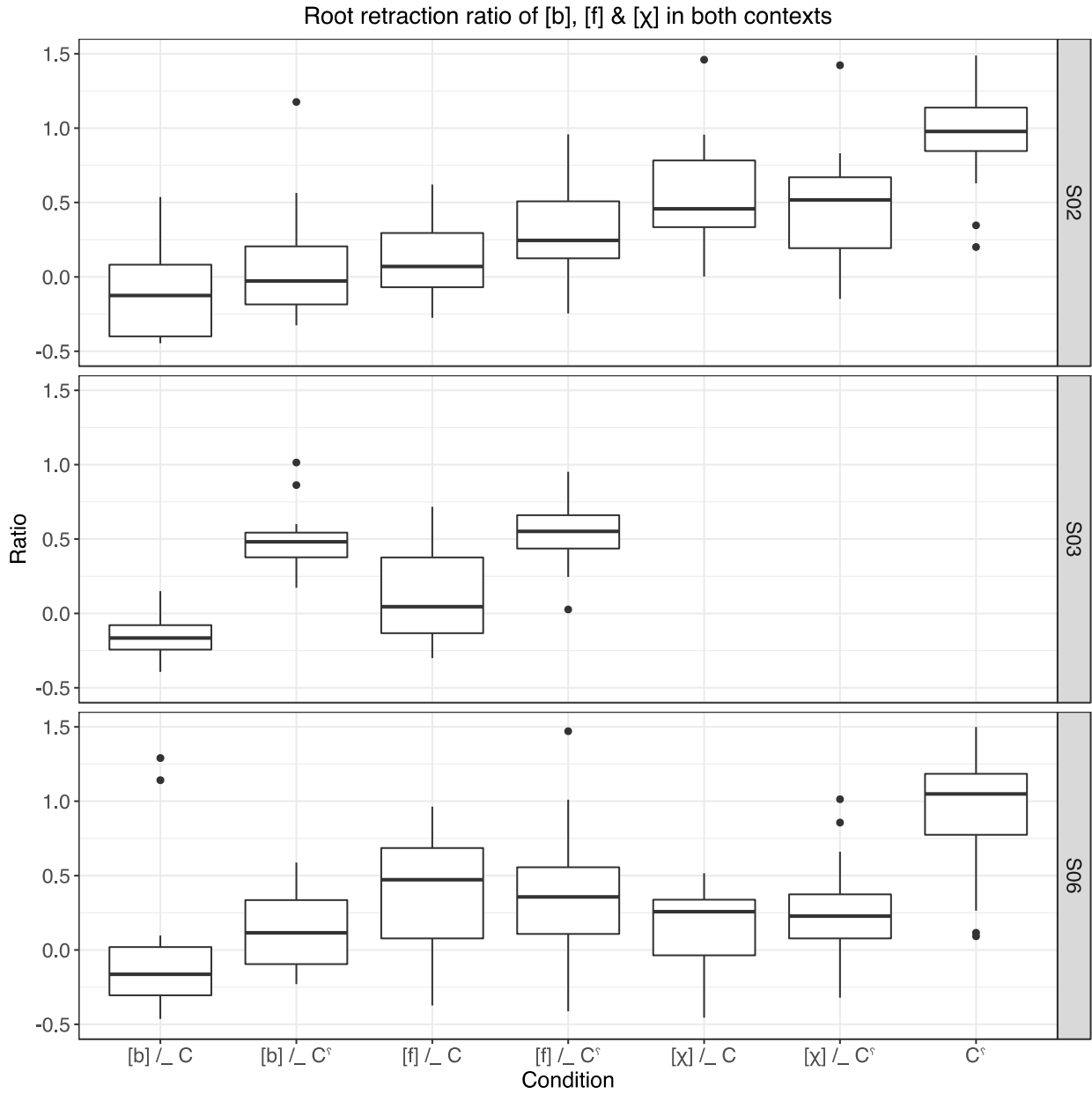


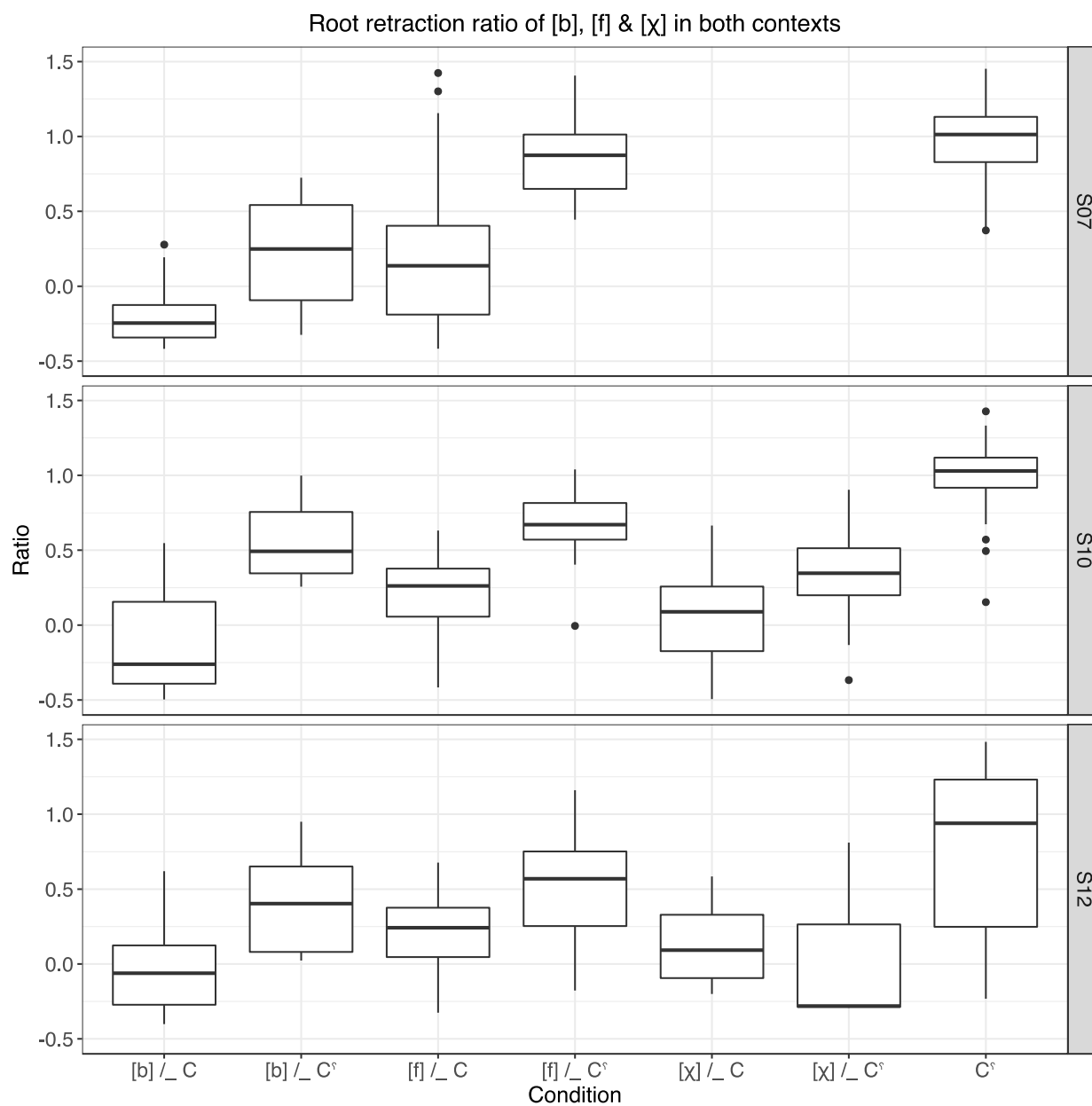


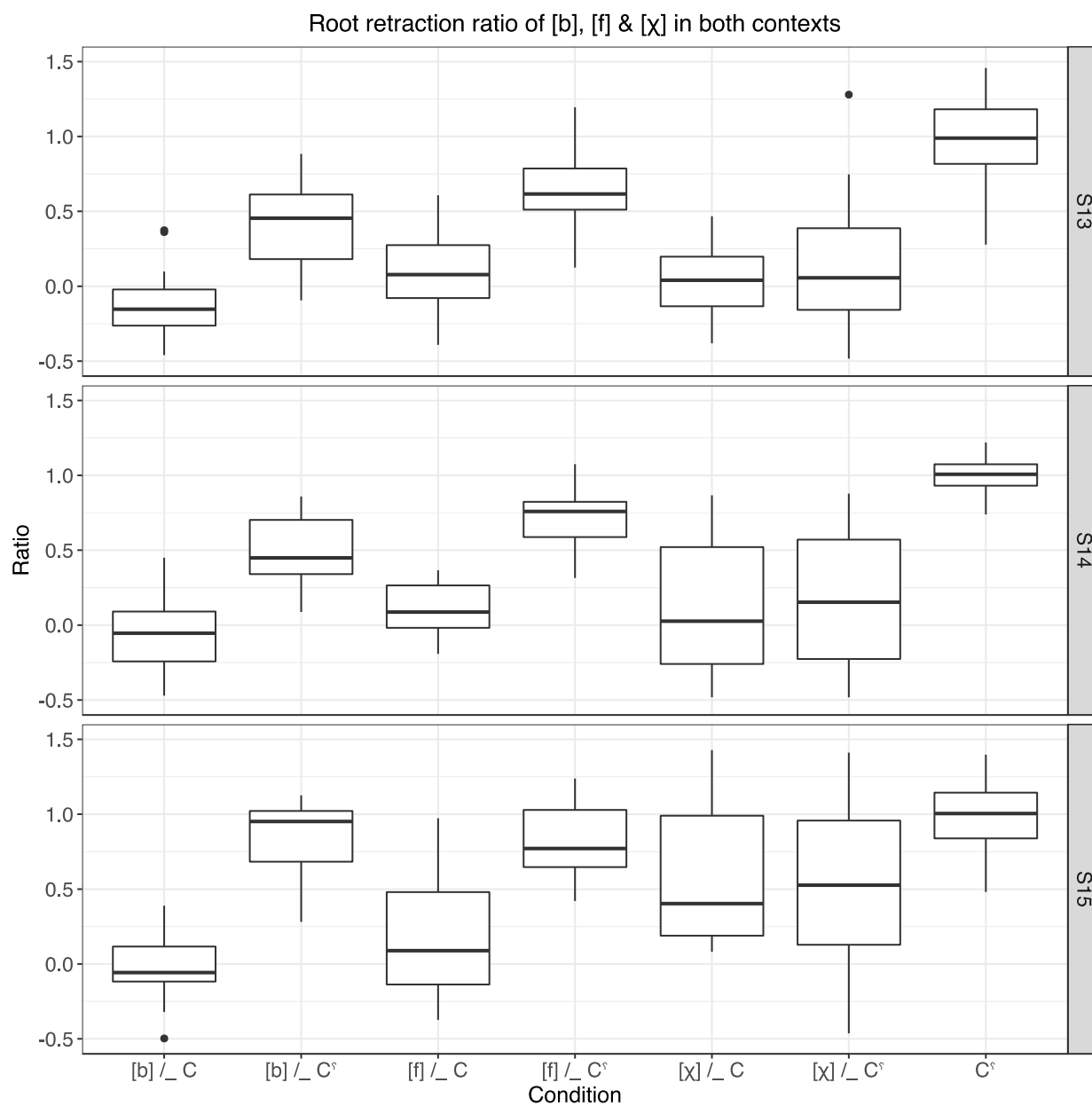


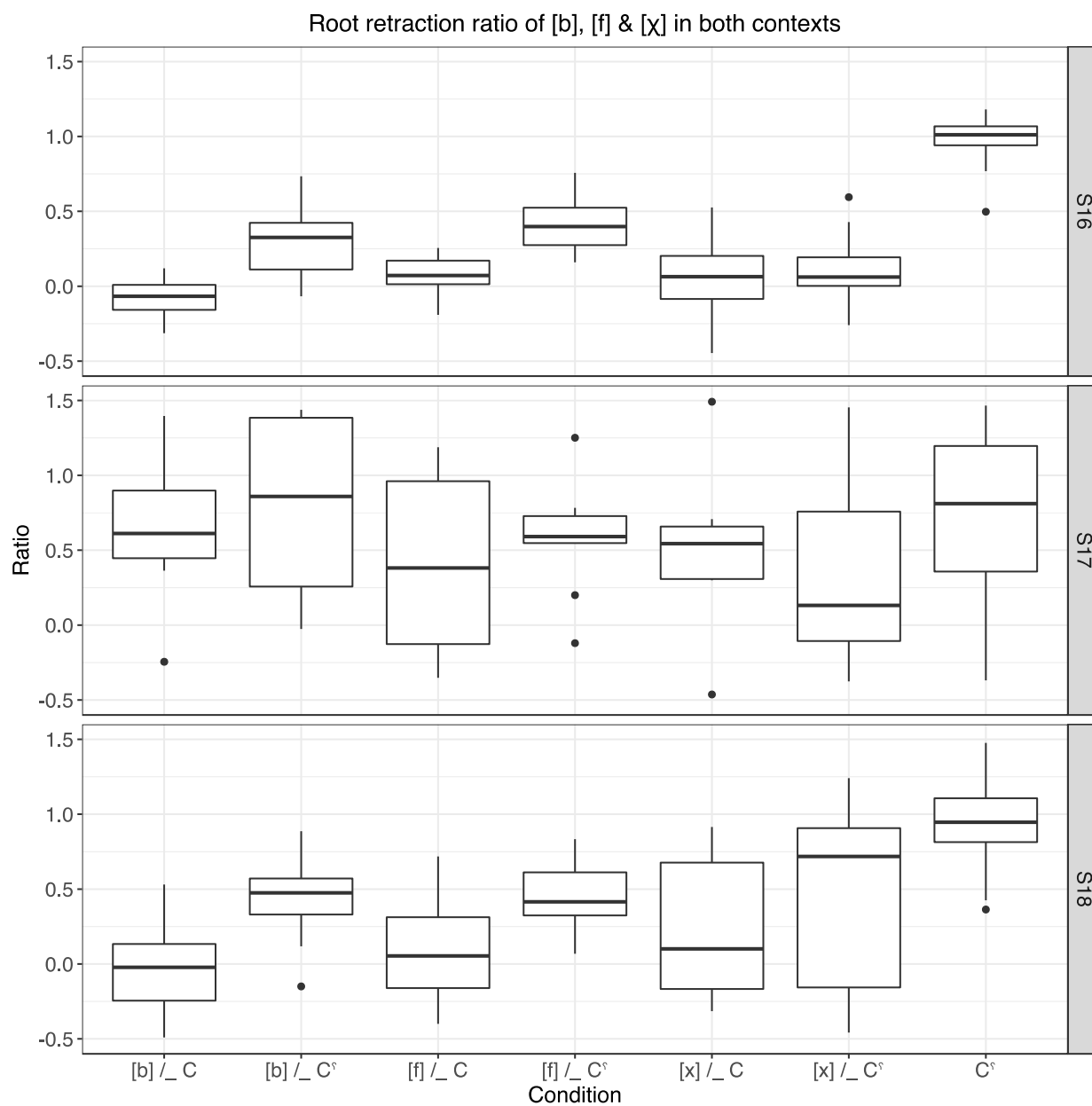


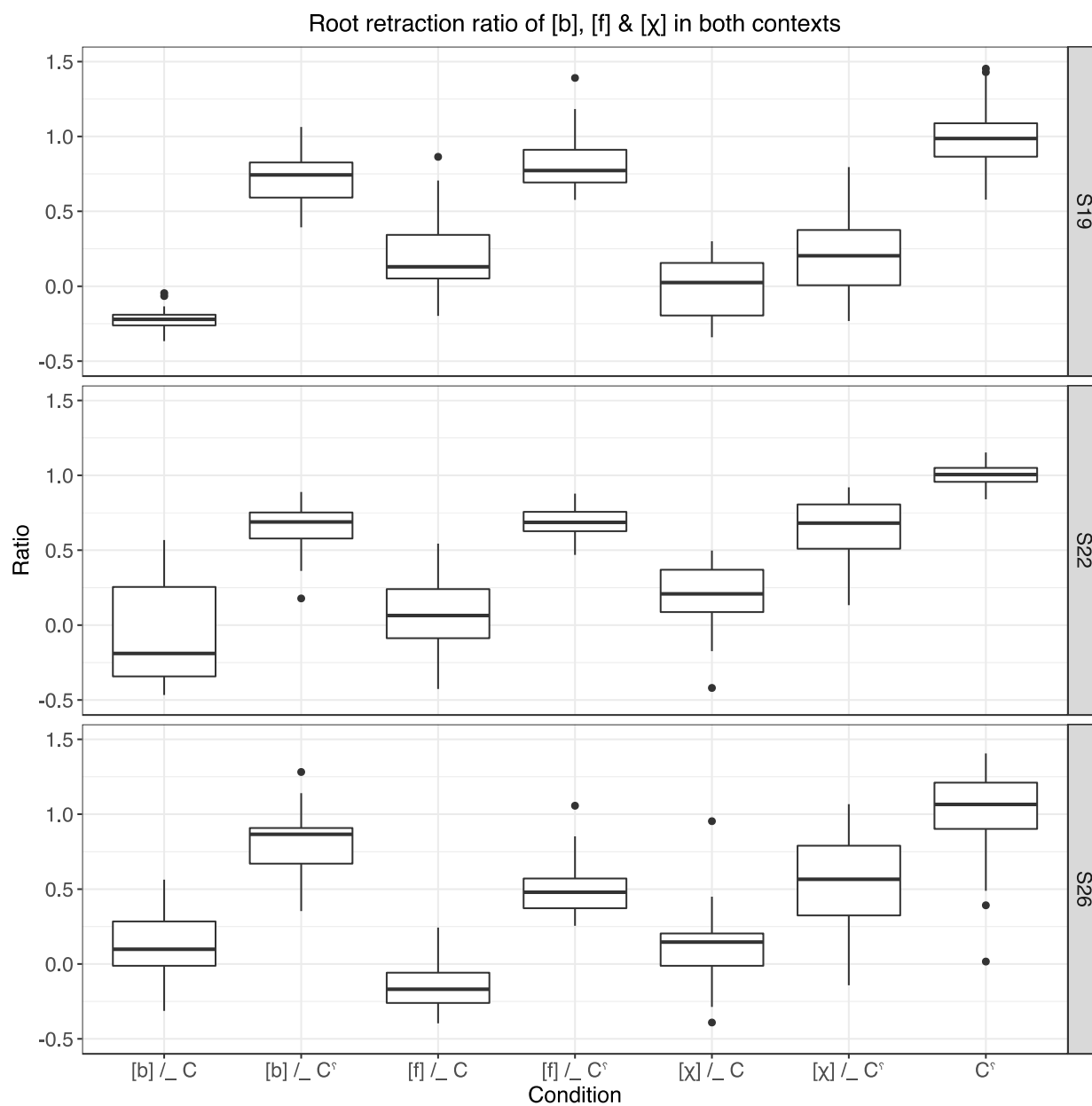
Appendix 3: Scatterplots of the root retraction ratio of [b], [f] and [χ]



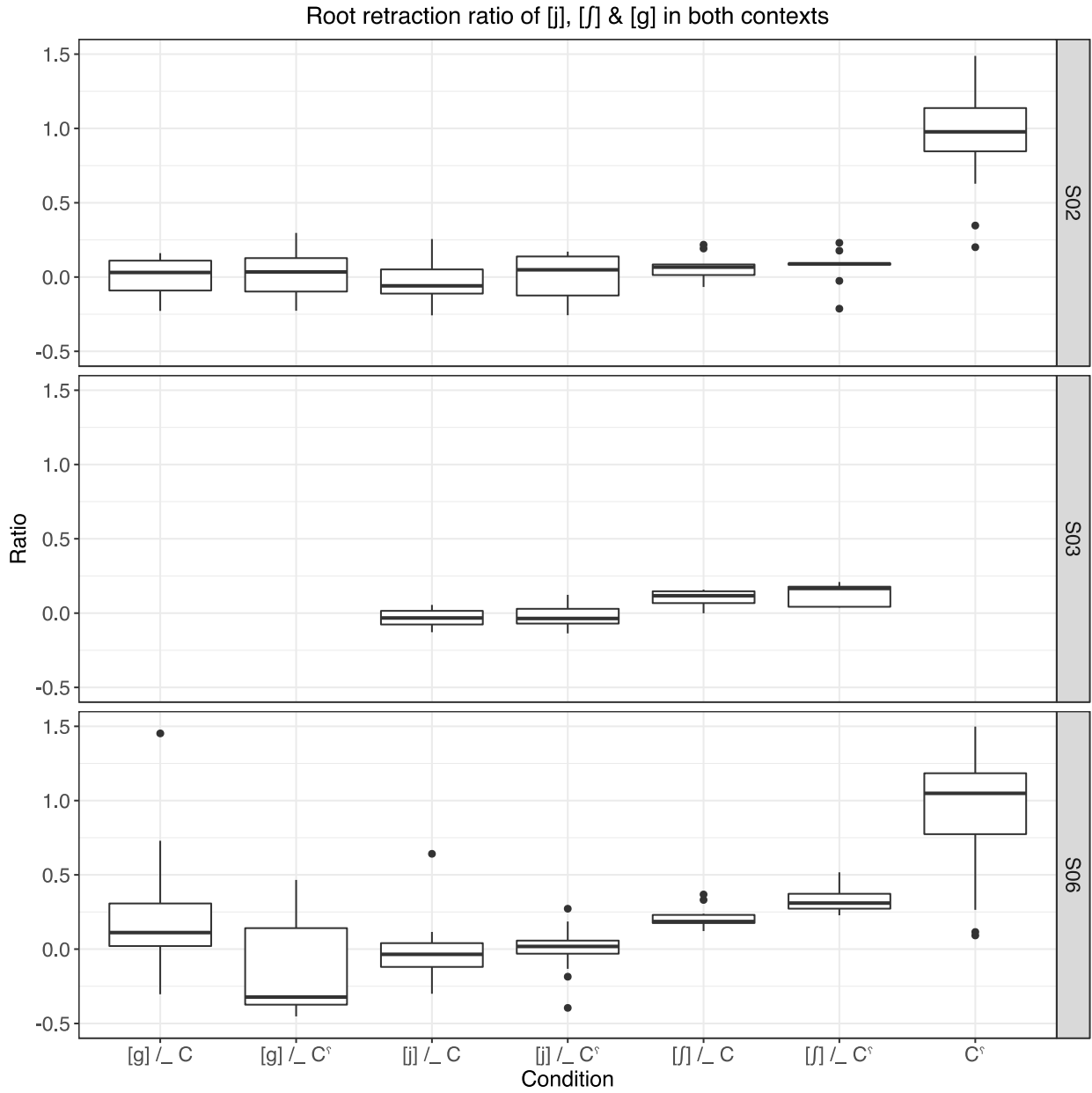


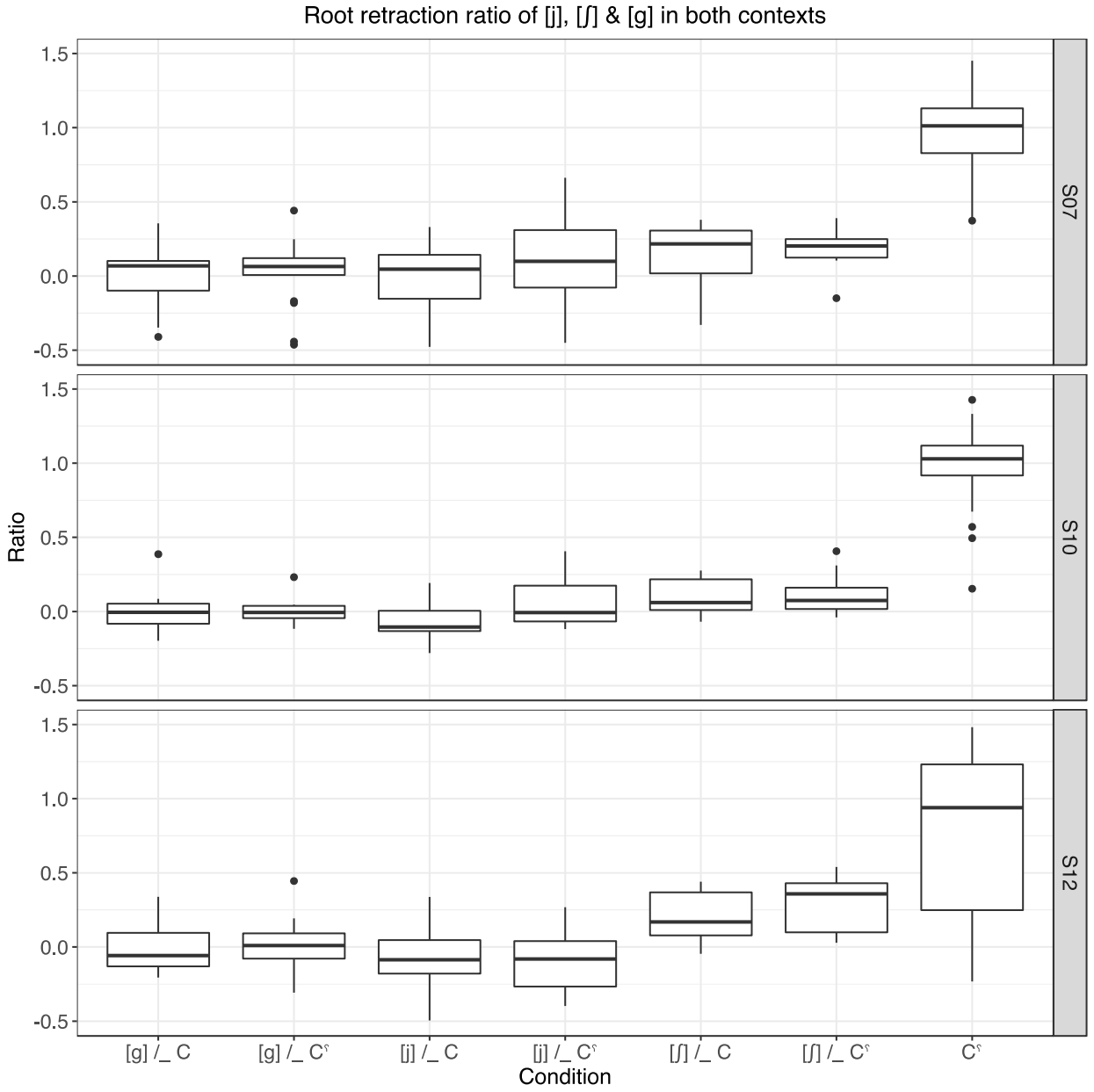


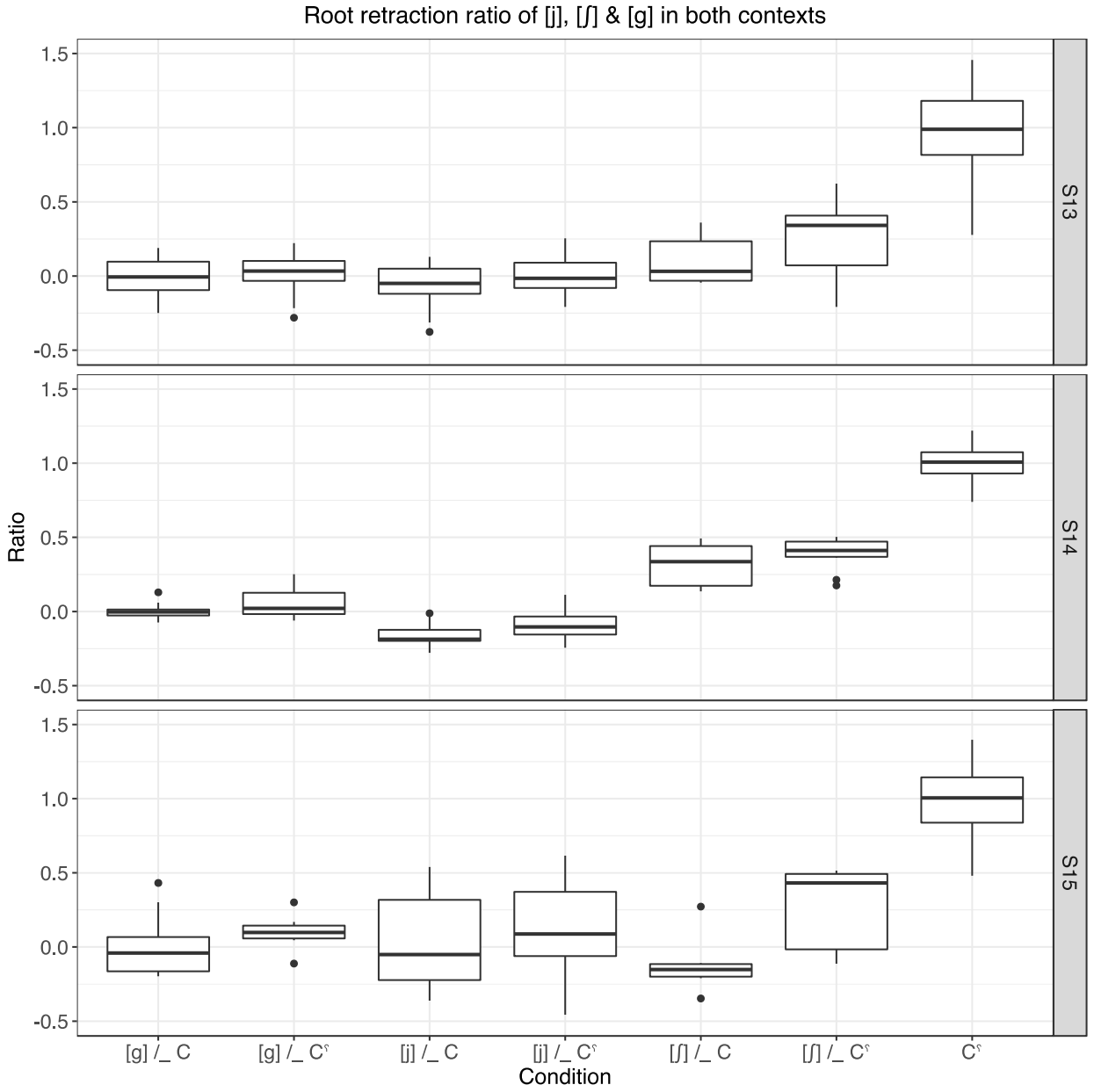


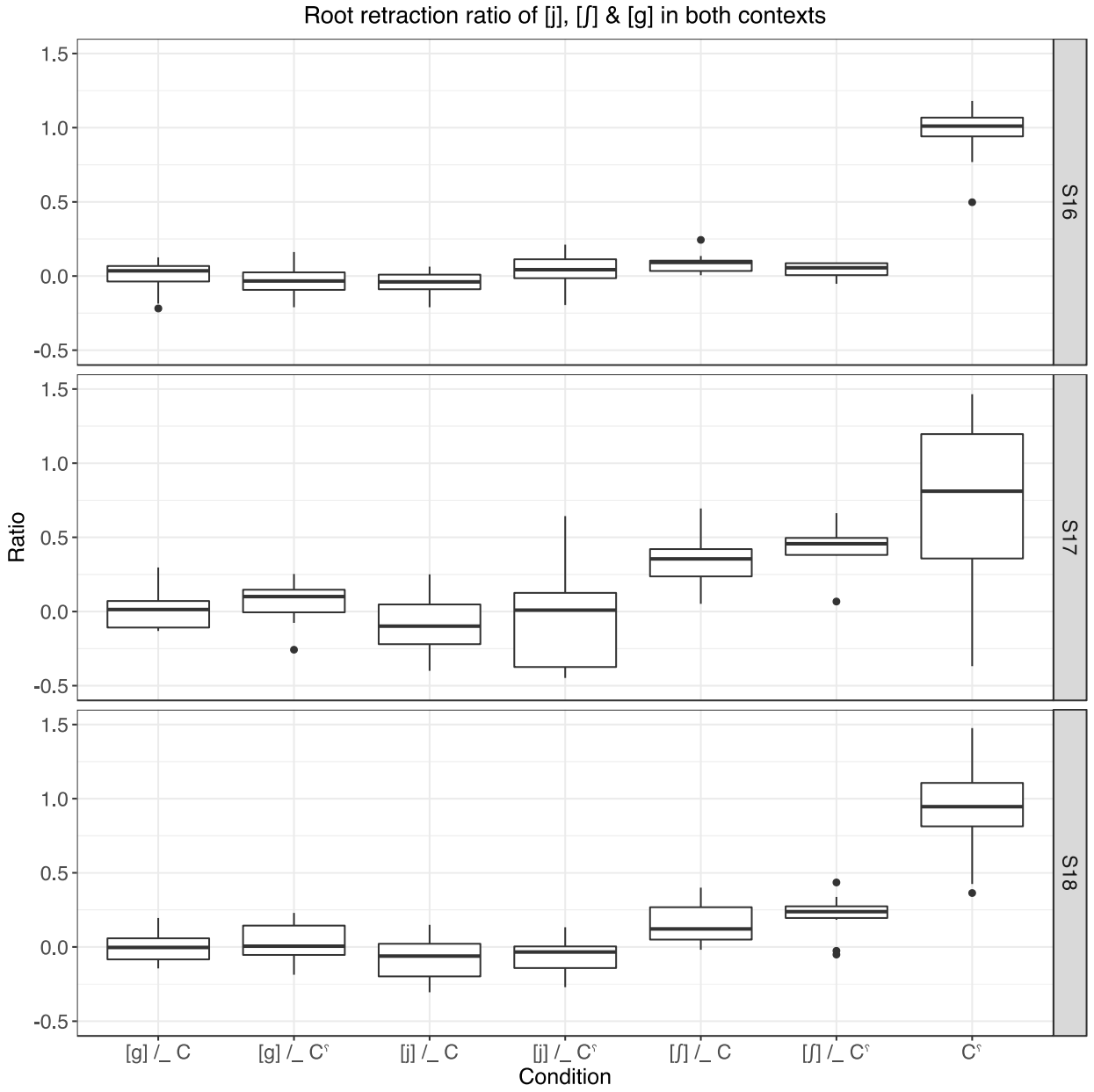


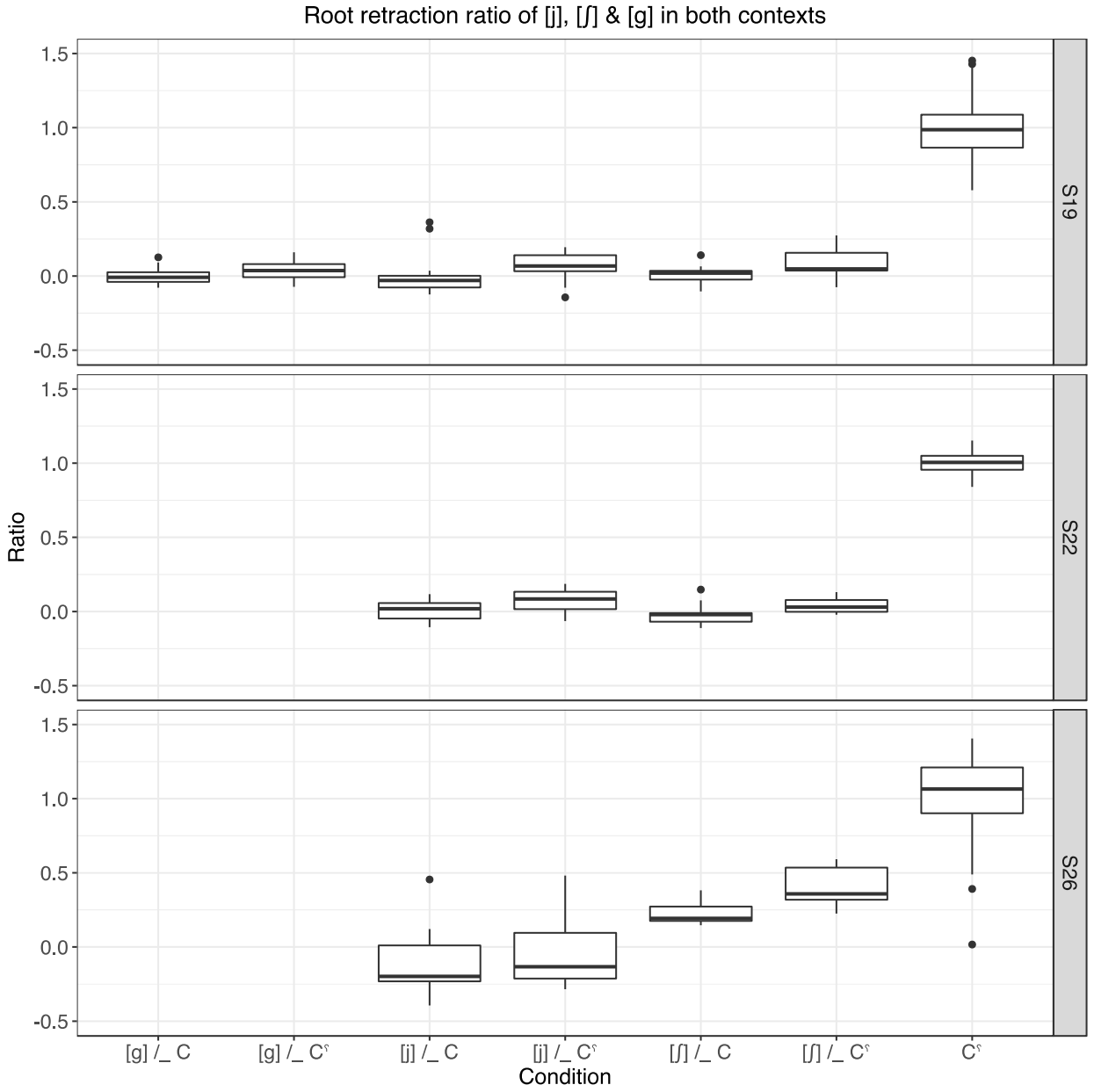
Appendix 4: Scatterplots of the root retraction ratio of [j], [ʃ] and [g]



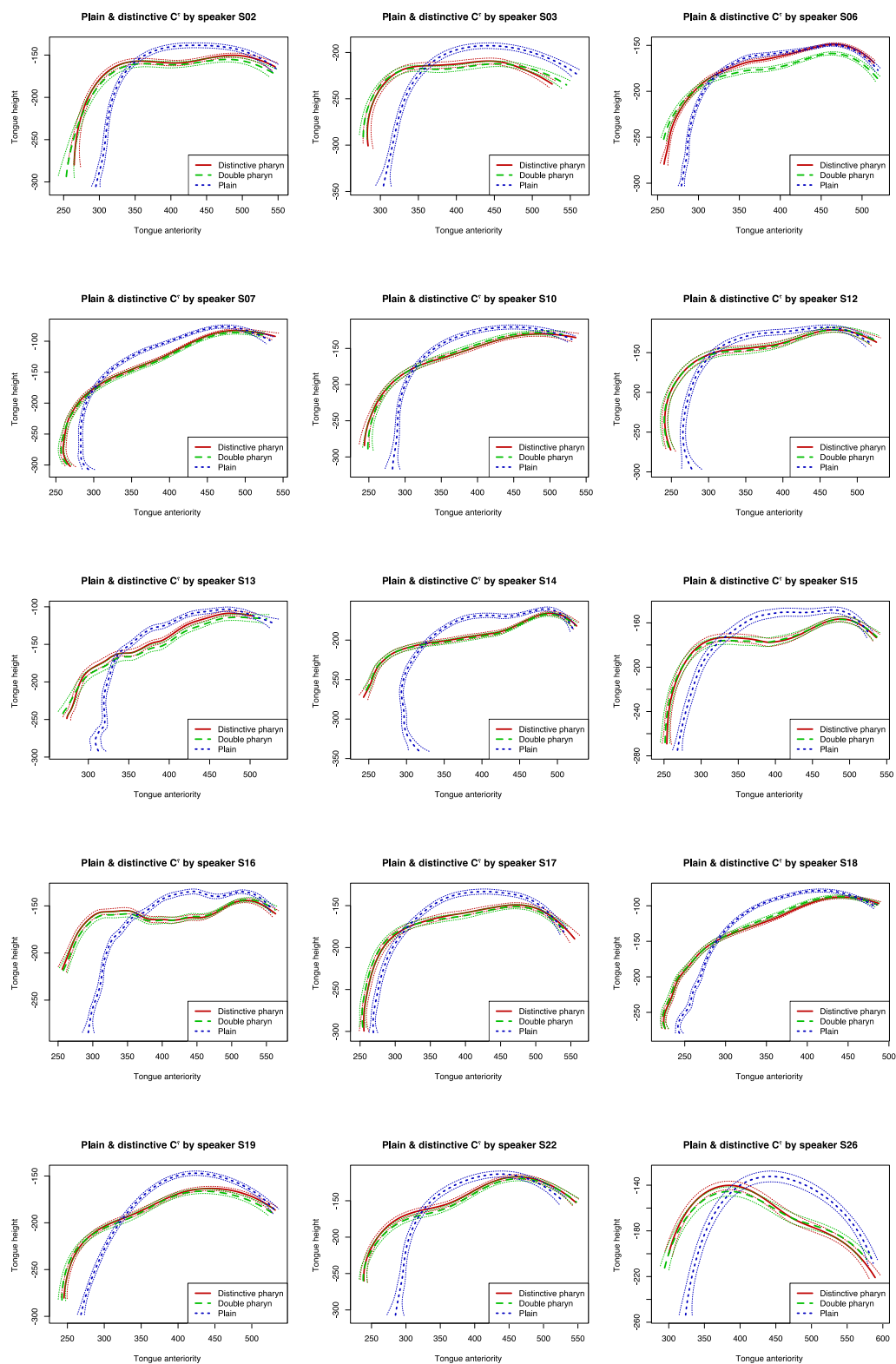




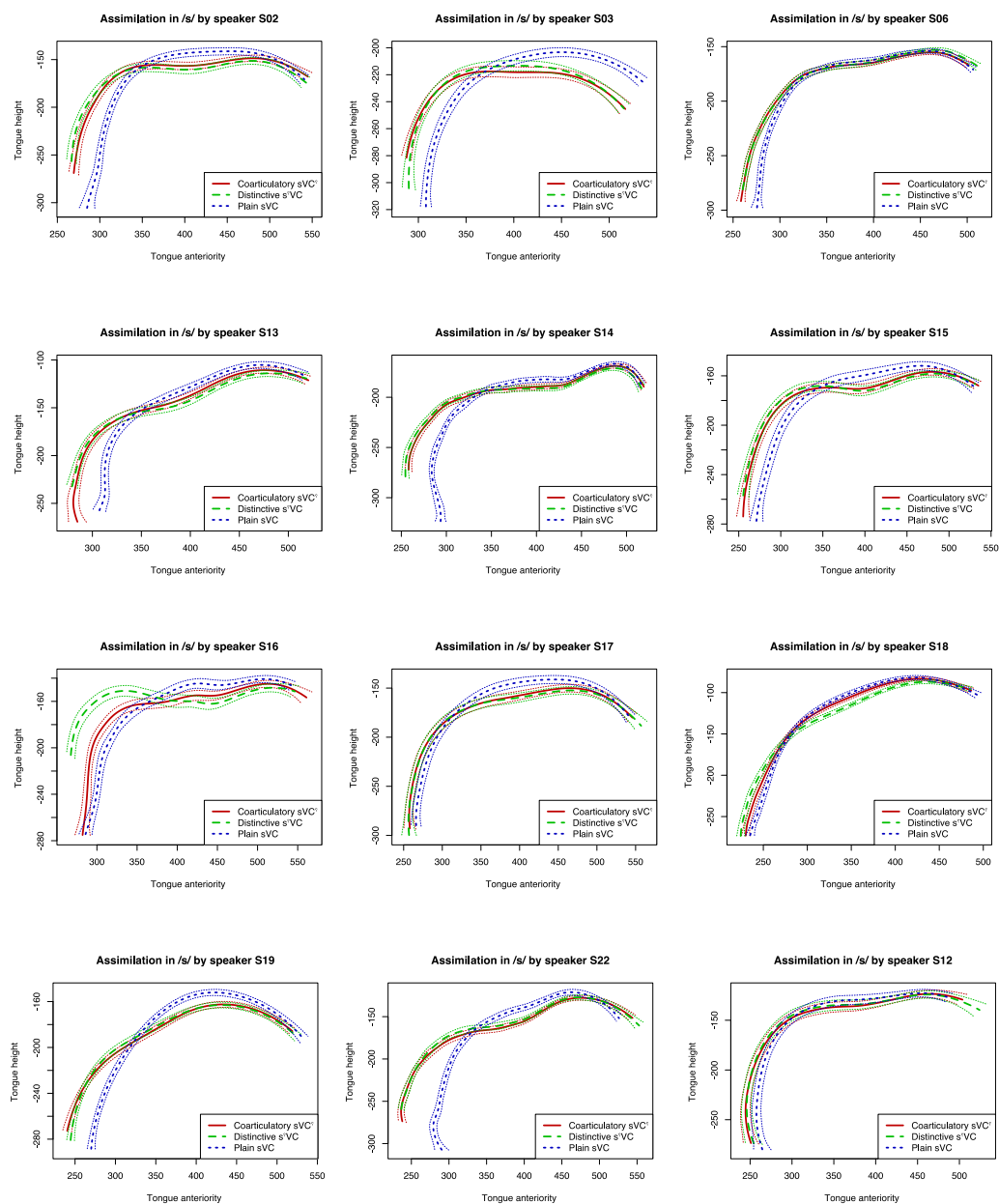




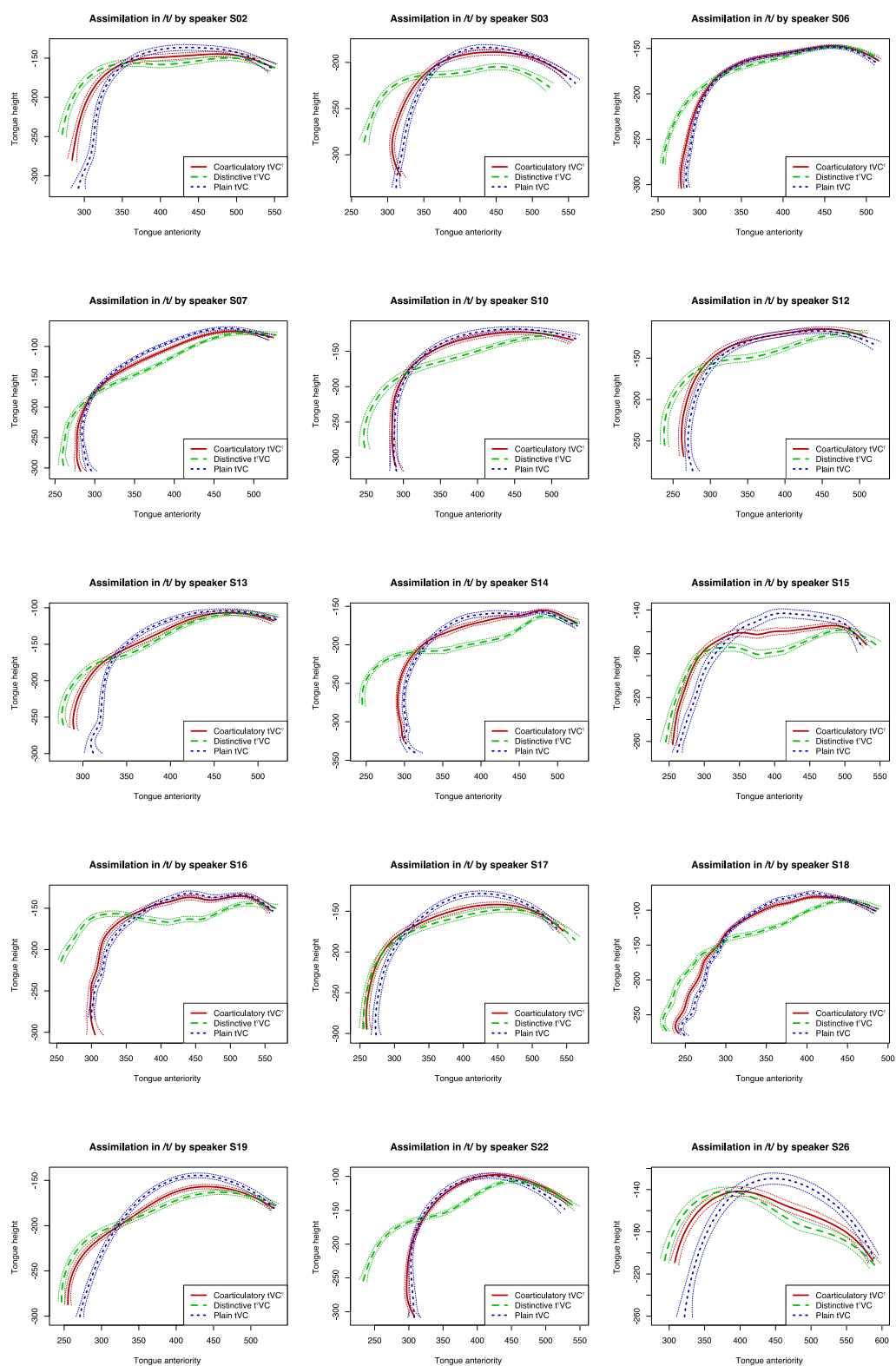
Appendix 5: Individual SS-ANOVA plots of plain C and distinctive C^s (Figure 3.9)



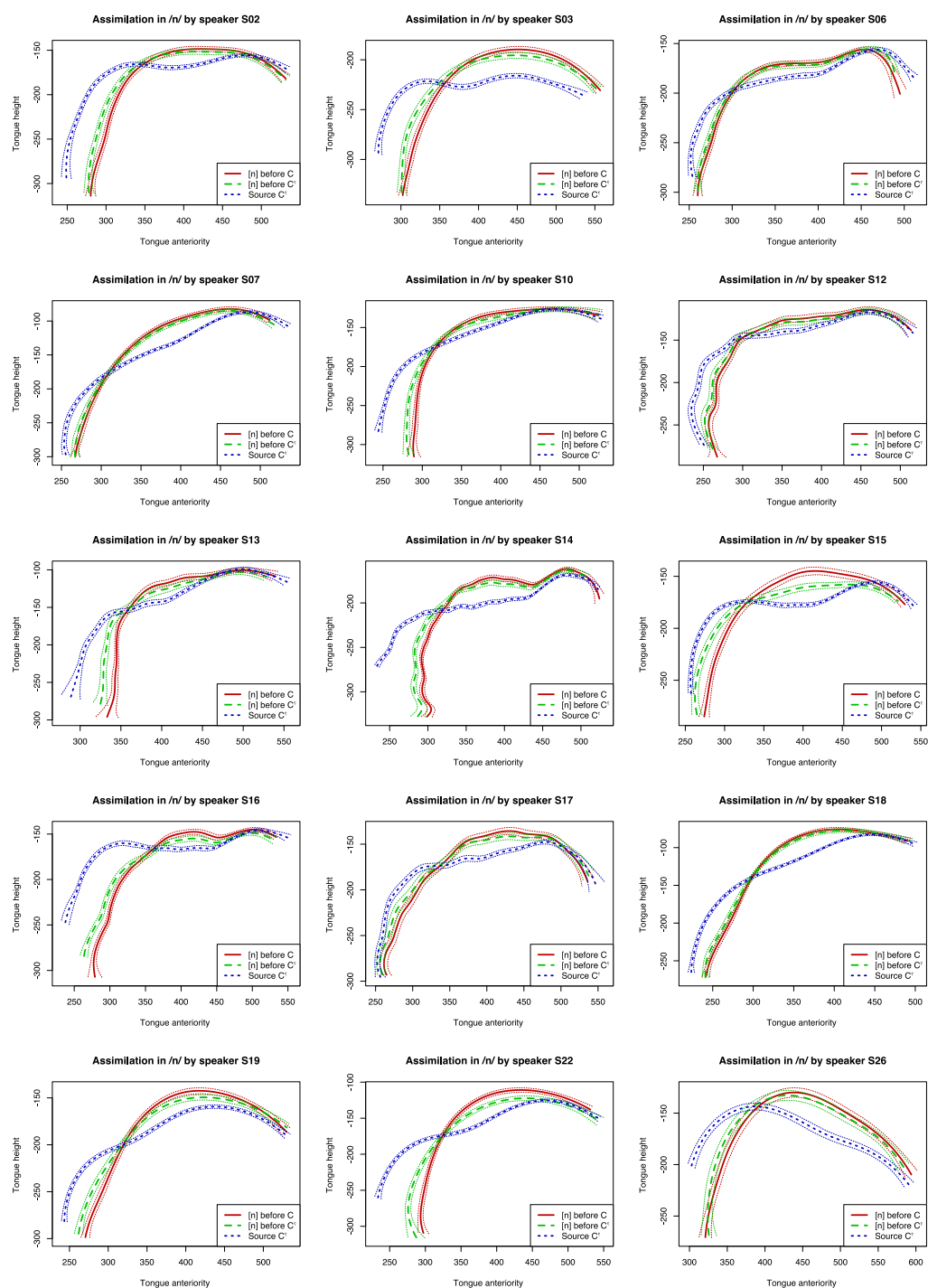
Appendix 6: Individual SS-ANOVA plots of /s/ in both contexts (Figure 3.14)



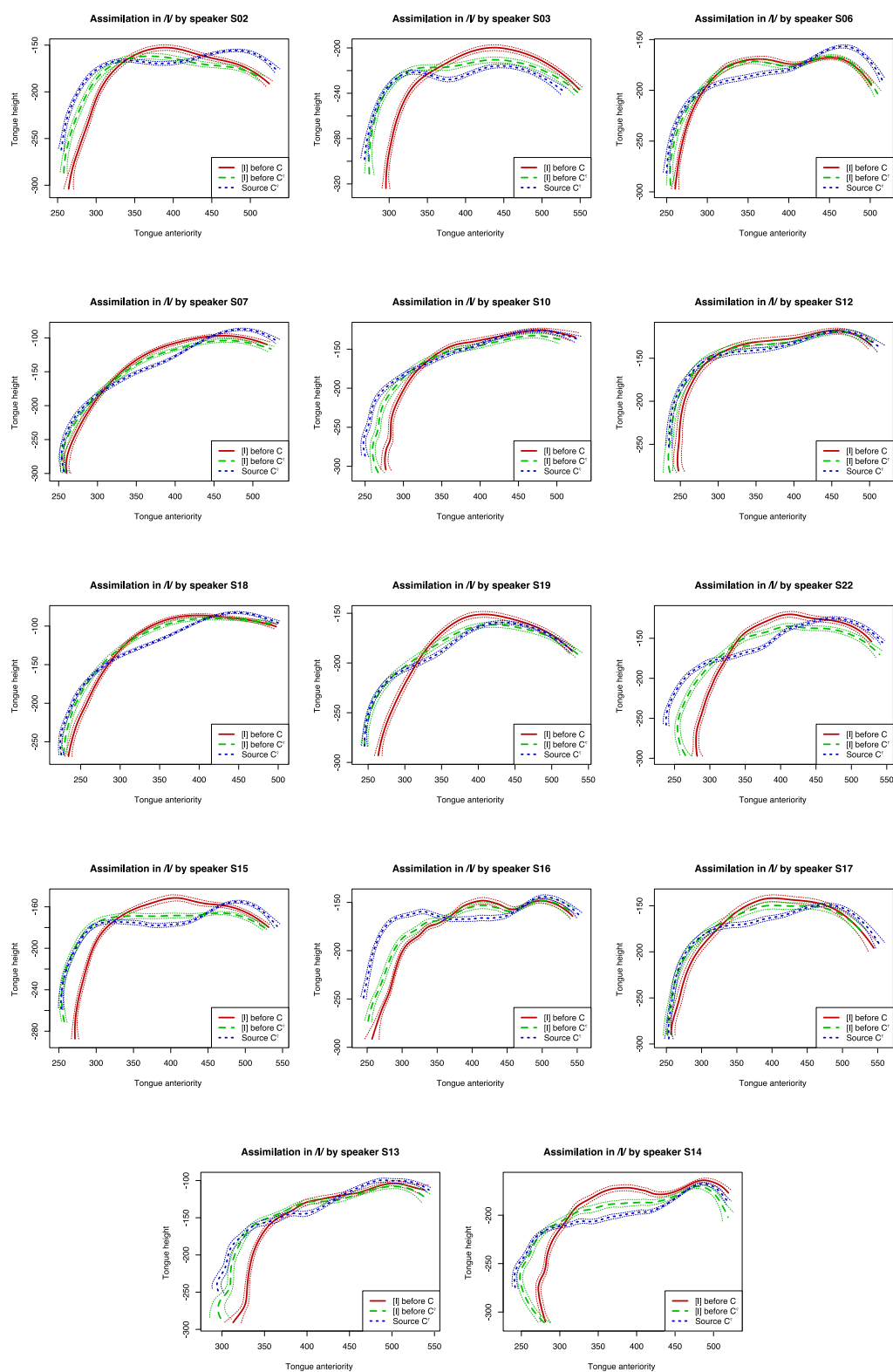
Appendix 7: Individual SS-ANOVA plots of /t/ in both contexts (Figure 3.17)



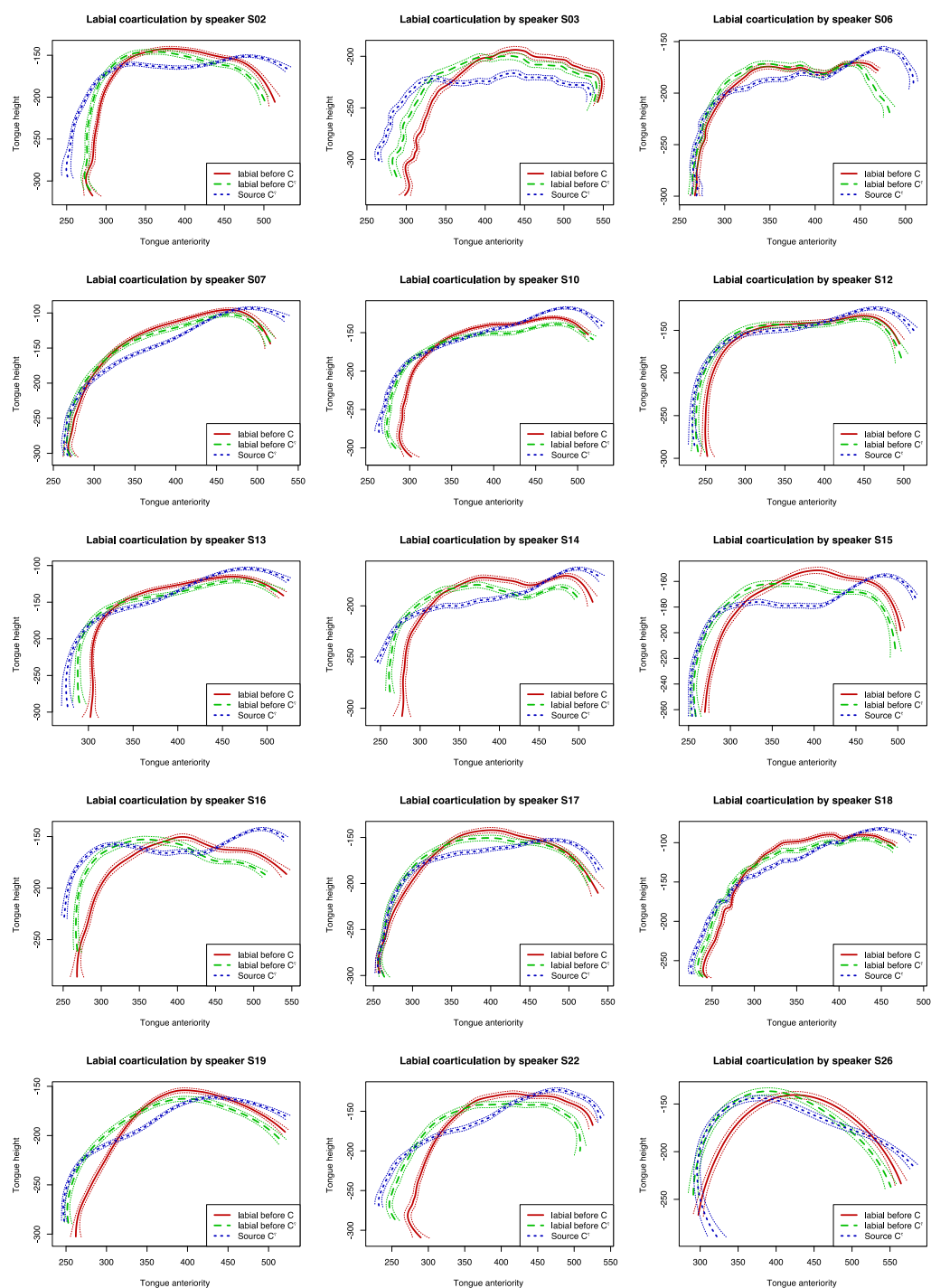
Appendix 8: Individual SS-ANOVA plots of /n/ in both contexts (Figure 4.2)



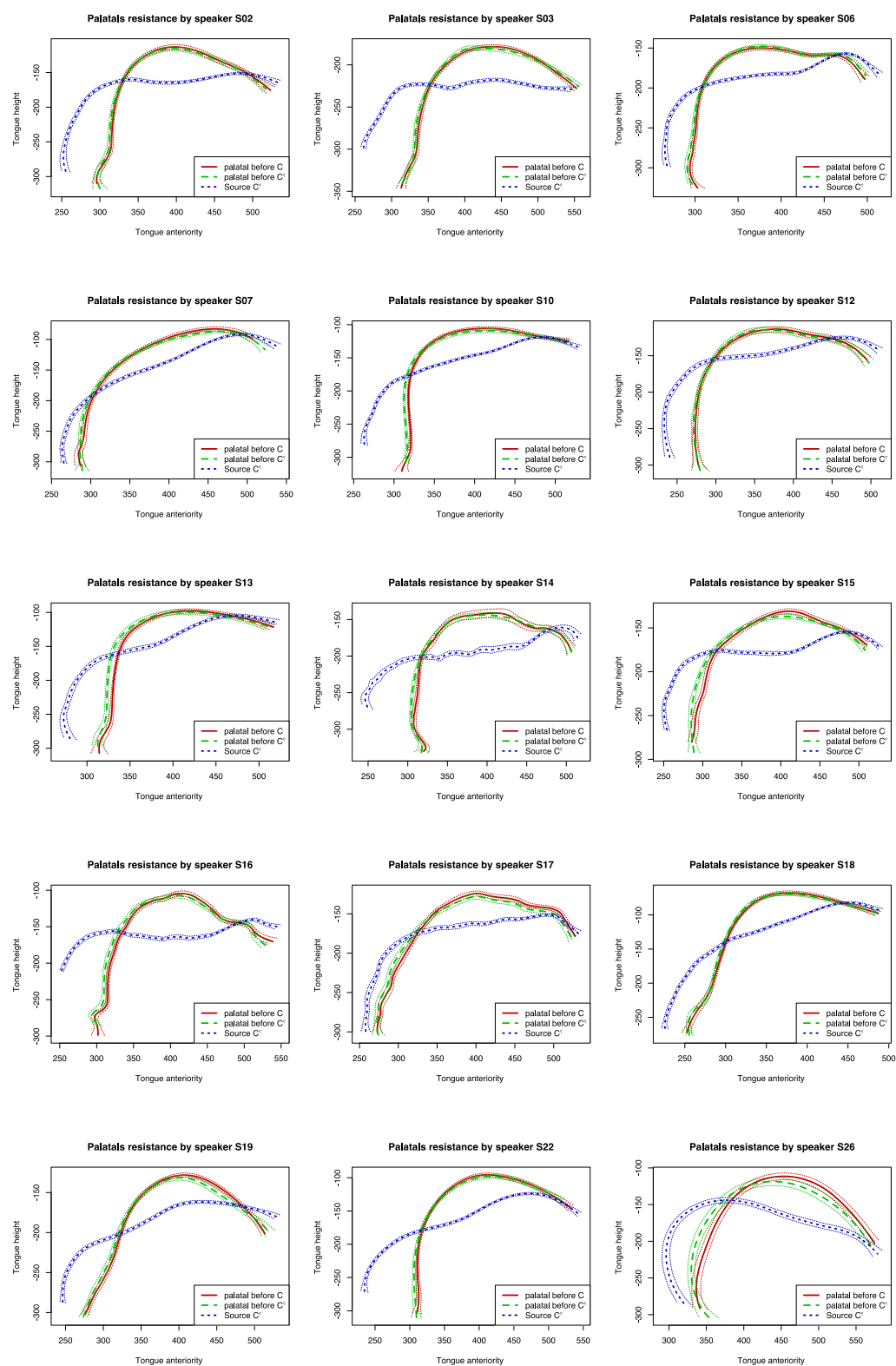
Appendix 9: Individual SS-ANOVA plots of /l/ in both contexts (Figure 4.5)



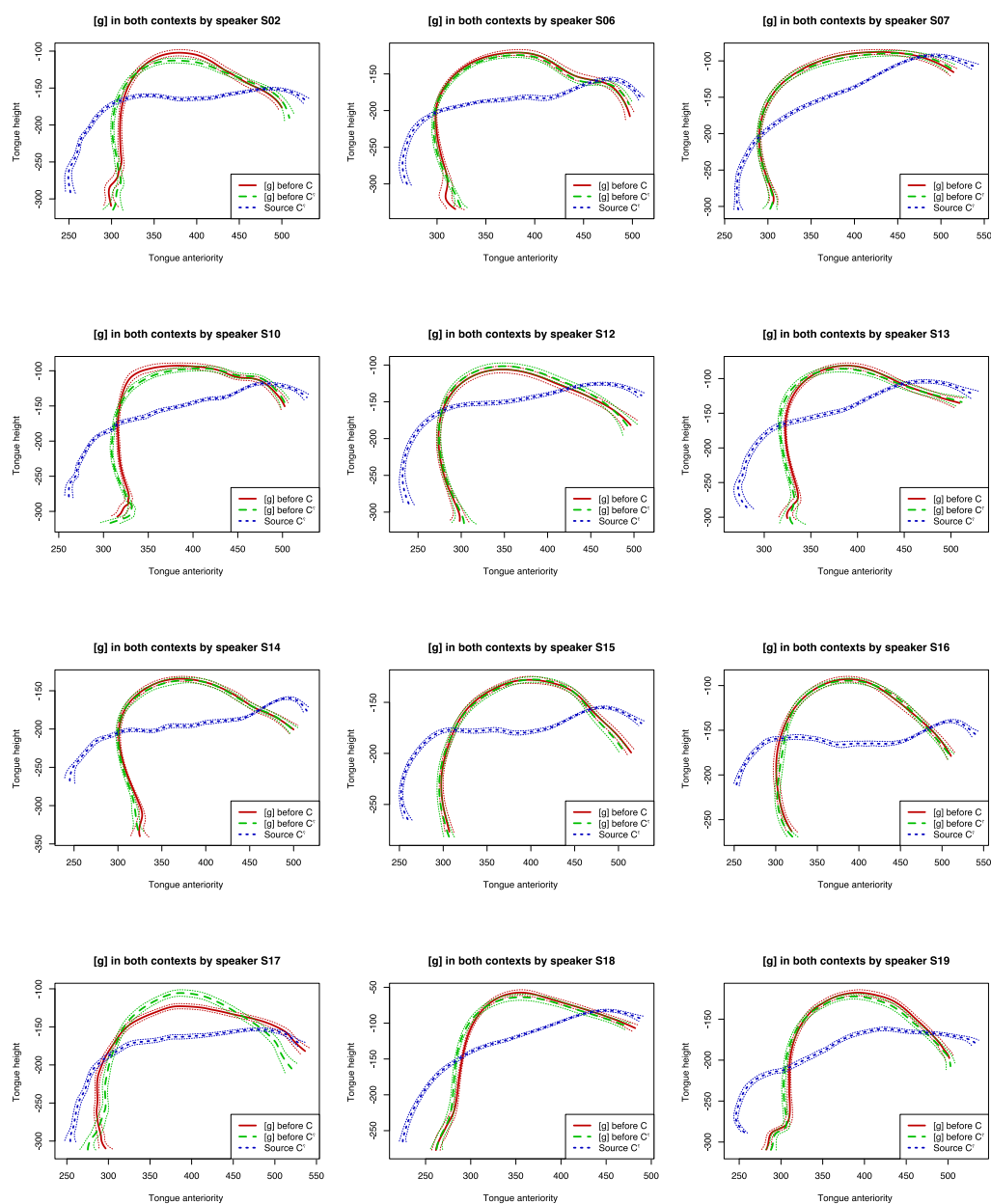
Appendix 10: Individual SS-ANOVA plots of labials in both contexts (Figure 5.6)



Appendix 11: Individual SS-ANOVA plots of palatals in both contexts (Figure 5.12)



Appendix 12: Individual SS-ANOVA plots of velar [g] in both contexts (Figure 5.20)



Appendix 13: Individual SS-ANOVA plots of uvular [χ] in both contexts (Figure 5.24)

